Estimation and Inference in Univariate and Multivariate Log-GARCH-X Models When the Conditional Density is Unknown

Genaro Sucarrat and Steffen Grønneberg and Alvaro Escribano

BI Norwegian Business School, BI Norwegian Business School, Universidad Carlos III de Madrid

11. August 2013

Estimation and Inference in Univariate and Multivariate Log-GARCH-X Models When the Conditional Density is Unknown *

Genaro Sucarrat†, Steffen Grønneberg‡and Álvaro Escribano§

First version: 9 June 2010
This version: 23rd February 2015

Abstract

Exponential models of Autoregressive Conditional Heteroscedasticity (ARCH) are of special interest, since they enable richer dynamics (e.g. contrarian or cyclical), provide greater robustness to jumps and outliers, and guarantee the positivity of volatility. The latter is not guaranteed in ordinary ARCH models, in particular when additional exogenous and/or predetermined variables ("X") are included in the volatility specification. We propose a general framework for the estimation and inference in univariate and multivariate Generalised log-ARCH-X (i.e. log-GARCH-X) models when the conditional density is not known. The framework employs (V)ARMA-X representations and relies on a bias-adjustment in the log-volatility intercept. The bias is induced by (V)ARMA estimators, but the remaining parameters are consistently estimated by (V)ARMA methods. We derive a simple formula for the bias-adjustment, and a closed-form expression for its asymptotic variance. Next, we show that adding exogenous or predetermined variables and/or increasing the dimension of the model does not change the structure of the problem. Accordingly, the univariate bias-adjustment is applicable not only in univariate log-GARCH-X models, but also in multivariate log-GARCH-X models. An empirical application illustrates the usefulness of the methods.

JEL Classification: C22, C32, C51, C52
Keywords: Log-GARCH-X, ARMA-X, multivariate log-GARCH-X, VARMA-X

*An earlier version of this paper was entitled "The Power Log-GARCH Model", see Sucarrat and Escribano (2010). We are grateful to Jonas Andersson, Luc Bauwens, Christian Francq, Andrew Harvey, Emma Iglesias, Sebastien Laurent, Enrique Sentana and seminar and conference participants at Nuffield College (Oxford University), Université de Lille, Universidad Carlos III de Madrid, BI Norwegian Business School (Oslo), IHS (Vienna), 2nd. Rimini Time Series Workshop (2013), 21st. SNDE Symposium (Milan, 2013), ESEM 2011 (Oslo), Interdisciplinary Workshop in Louvain-la-Neuve 2011, CFE conference 2010 (London), FIBE 2010 (Bergen), IWAP 2010 (Madrid) and Foro de Finanzas 2010 (Elche) for useful comments, suggestions and questions. Funding from the 6th. European Community Framework Programme, MICIN ECO2009-08308 and from The Bank of Spain Excellence Program is gratefully acknowledged.

†Corresponding author. Department of Economics, BI Norwegian Business School, Nydalsveien 37, 0484 Oslo, Norway. Email genaro.sucarrat@bi.no, phone +47+46410779, fax +47+23264788. Webpage: http://www.sucarrat.net/

‡Department of Economics, BI Norwegian Business School. Email: steffeng@gmail.com.

§Department of Economics, Universidad Carlos III de Madrid (Spain). Email: alvaroe@eco.uc3m.es

1 Introduction

The Autoregressive Conditional Heteroscedasticity (ARCH) class of models due to Engle (1982) is useful in a wide range of applications. In finance in particular, it has been extensively used to model the clustering of large (in absolute value) financial returns. Engle (1982) himself, however, originally motivated the class as useful in modelling the time-varying conditional uncertainty (i.e. conditional variance) of economic variables in general, and of UK inflation in particular. Other areas of application include, amongst other, the uncertainty of electricity prices (e.g. Koopman et al. (2007)), the evolution of temperature data (e.g. Franses et al. (2001)) and – more generally – positively valued variables, i.e. socalled Multiplicative Error Models (MEMs), see Brownlees et al. (2012).

Within the ARCH class of models exponential versions are of special interest. This is because they enable richer autoregressive volatility dynamics (e.g. contrarian or cyclical) compared with non-exponential ARCH models, and because their fitted values of volatility are guaranteed to be positive. The latter is not necessarily the case for ordinary (i.e. non-exponential) ARCH models, in particular when additional exogenous or predetermined variables ("X") are included in the volatility equation.

In fact, the greater the dimension of X, the more restrictions are needed in order to ensure positivity. Another desirable property is that volatility forecasts are more robust to jumps and outliers. Robustness can be important in order to avoid volatility forecast failure subsequent to jumps and outliers.

The log-GARCH class of models can be viewed as a dynamic version of Harvey's (1976) multiplicative heteroscedasticity model, and was first proposed independently by Pantula (1986), Geweke (1986) and Milhøj (1987). Engle and Bollerslev (1986) argued against log-ARCH models because of the possibility of applying the log-operator (in the log-ARCH terms) on zero-values, which occurs whenever the error term in a regression equals zero. A solution to this problem, however, is provided in Sucarrat and Escribano (2013) for the case where the zero-probability is zero (e.g. because zeros are due to discreteness or missing values).[1] The solution is only available when estimation is via the (V)ARMA representation. Finally, two competing classes of exponential ARCH models are Nelson's (1991) EGARCH and Harvey's (2013) Beta-t-EGARCH model. The former has proved to be much more difficult theoretically (more on this below), and the latter is not – by its very nature – amenable to the assumption of an unknown conditional density (i.e. the conditional density must be known).

The assumption that the conditional density is unknown is particularly convenient from a practitioner's point of view, since the user then does not need to worry about changing the conditional density from application to application, or alternatively to work with a sufficiently general density that will often make estimation and inference numerically more challenging. This explains the attraction of Quasi Maximum Likelihood Estimators (QMLEs). In the univariate case consistency and asymptotic normality of QMLE for GARCH models under mild conditions were first established by Berkes et al. (2003) and Francq and Zakoïan (2004). In the exponential case most of the attention has been directed at Nelson's (1991) EGARCH, whose asymptotic properties have turned out to be very difficult to establish, see e.g. Straumann and Mikosch (2006). Only recently was consistency and asymptotic normality proved (for the univariate EGARCH(1,1) only) under the complicated condition of continuous invertibility, see Wintenberger (2013). The log-GARCH model is much more tractable. Francq et al. (2013) prove consistency and asymptotic normality of the Gaussian QMLE for an asymmetric log-GARCH(p, q) model under mild conditions. Their method does not employ ARMA representations, which means it is more efficient when the conditional error is normal or close to normal, but not when the conditional density is fat-tailed, see the asymptotic efficiency comparison in Francq and Sucarrat (2013)). Moreover, the estimator of Francq et al. (2013) cannot handle zero-errors or missing values as suggested in Sucarrat and Escribano (2013). Finally, Francq and Sucarrat (2013) propose an estimator that achieves efficiency for conditional densities that are normal or close to the normal, by combining the ARMA-approach with the Centred Exponential Chi-Squared as instrumental QML-density. In the multivariate case, QML results have been established for the BEKK model of Engle and Kroner (1995) by Comte and Lieberman (2003), for an ARMA-GARCH with constant conditional correlations (CCCs) by Ling and McAleer (2003), for a fac-

[1]The same idea can be extended to the case where the zero-probability is non-zero and time-varying.

tor GARCH model by Hafner and Preminger (2009), for a multivariate GARCH with CCCs by Francq and Zakoïan (2010) and for a multivariate GARCH with stochastic correlations by Francq and Zakoïan (2014) under the assumption that the system is estimable equation-by-equation.[2] For exponential ARCH models there are no multivariate results. Kawakatsu (2006) has proposed a multivariate exponential ARCH model, the matrix exponential GARCH, which contains a multivariate version of Nelson's 1991 model. But there are no proofs for the estimation and inference methods that he proposes.

This paper makes four contributions. It is well-known that all the coefficients apart from the log-volatility intercept in a univariate log-GARCH specification can be estimated consistently (under suitable assumptions) via an ARMA representation, see for example Psaradakis and Tzavalis (1999), and Francq and Zakoïan (2006). However, the estimate of the log-volatility intercept will be asymptotically biased, and the bias is made up of a log-moment expression that depends on the unknown density of the conditional error. We propose a simple estimator of the log-moment expression that is made up of the empirical residuals of the ARMA regression, and derive an expression for its asymptotic variance (Sections 2.3-2.4). The practical consequence of this is that the log-volatility intercept can be estimated consistently, and hence that *all* the log-GARCH parameters can be estimated consistently via the ARMA representation.

In the second contribution of our paper (Section 2.5), we show that the addition of exogenous, determinstic and/or predetermined conditioning variables, i.e. the log-GARCH-X model, does not alter the relation between the ARMA coefficients and the log-GARCH coefficients. So consistent estimation of the ARMA-X representation will produce exactly the same bias as earlier, and the bias correction procedure described above is applicable also for ARMA-X models.

In the third contribution (Section 3) we propose a multivariate log-GARCH-X model that admits time-varying conditional correlations. The model has a VARMA-X representation with a vector of error-terms. The vector is either IID, which corresponds to the Constant Conditional Correlation (CCC) case, or independent but non-identical (ID), which corresponds to the time-varying correlations case. In both cases, however, each entry in the vector of errors is marginally IID. So the bias-correction from the univariate case can be used equation-by-equation – under suitable assumptions – subsequent to the estimation of the VARMA-X representation.

In the fourth contribution (Section 4) we illustrate the usefulness of our results by an application to the modelling of the uncertainty of electricity prices. Electricity prices are characterised by autoregressive persistence, day-of-the week effects, large spikes or jumps, ARCH and non-normal conditional errors that are possibly skewed. For robust (to jumps) forecasts of uncertainty (i.e. volatility) that accommodates all these characteristics, the log-GARCH-X model is particularly suited. The investigation shows that volatility can be substantially underestimated if sufficient ARCH-lags and day-of-the-week effects are not accommodated.

The rest of the paper is organised as follows. The next section, section 2, presents

[2] Jeantheau (1998) established general conditions for strong consistency for QML estimation of multivariate GARCH models. However, as pointed out by Ling and McAleer (2003, p. 281), his results are based on the unrealistic assumption that the initial values are known.

the univariate log-GARCH model, the relation between the univariate log-GARCH model and its ARMA representation, and derives the log-moment estimator and its asymptotic variance. Also, it is shown that the addition of exogenous and predetermined variables does not alter the relationship between the log-GARCH and ARMA parameters. Section 3 shows how the ideas extend to the multivariate case. Section 4 contains our empirical application, whereas Section 5 concludes. Tables and Figures are placed at the end.

2 Univariate log-GARCH

2.1 Notation and specification

The univariate log-GARCH(p,q) model is given by

$$\begin{aligned} \epsilon_t &= \sigma_t z_t, \quad z_t \sim IID(0,1), \quad P(z_t = 0) = 0, \quad \sigma_t > 0, & (1) \\ \ln \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{p} \alpha_i \ln \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \ln \sigma_{t-j}^2, \quad t \in \mathbb{Z}, & (2) \end{aligned}$$

where p is the ARCH order and q is the GARCH order. In finance, ϵ_t is often interpreted as return or mean-corrected return, but more generally it is simply the error in a regression model. Throughout we will assume ϵ_t is observable and known. Of course, this is not a realistic nor a desirable assumption, but simply reflects the current state of the theoretical literature.[3] Denoting $p^* = \max\{p,q\}$, if the roots of the lag polynomial $1-(\alpha_1+\beta_1)L-\cdots-(\alpha_{p^*}+\beta_{p^*})L^{p^*}$ are all greater than 1 in modulus and if $|E(\ln z_t^2)| < \infty$, then $\ln \sigma_t^2$ is stable. For common densities like the Student's t with degrees of freedom greater than 2, and the Generalised Error Distribution (GED) with shape parameter greater than 1, then σ_t^2 will generally be stable as well if $\ln \sigma_t^2$ is stable. Practitioners are often interested in the dynamics of other powers than the 2nd., e.g. the 1st. power (i.e. the conditional standard deviation). For that purpose it should be noted that the dth. power log-GARCH(p,q) model can be written as

$$\ln \sigma_t^d = \alpha_{0,d} + \sum_{i=1}^{p} \alpha_i \ln |\epsilon_{t-i}|^d + \sum_{j=1}^{q} \beta_j \ln \sigma_{t-j}^d, \quad d > 0, \tag{3}$$

where $\alpha_{0,d} = \alpha_0 d/2$. This means a complete analysis of the dth. power log-GARCH model can be undertaken in terms of the $d = 2$ representation.

The log-GARCH model accommodates a broader range of persistency structures than the ordinary GARCH model. In particular, in contrast to the ordinary GARCH model, the unconditional autocorrelations of log-GARCH models depend on the distribution of z_t: The more fat-tailed, the weaker correlations. Also, the log-GARCH is capable of generating both weaker and stronger autocorrelations than the GARCH, and autocorrelation functions that decline either more rapidly or more slowly.

[3]To the best of our knowledge there are only two results in the literature that do not need to assume that ϵ_t is known, namely Ling and McAleer (2003) and Francq and Zakoïan (2004). Both accommodate the joint estimation of the mean and variance equations simultaneously.

2.2 The ARMA representation

If $|E(\ln z_t^2)| < \infty$, then the log-GARCH(p, q) model (1)-(2) admits the ARMA(p, q) representation

$$\ln \epsilon_t^2 = \phi_0 + \sum_{i=1}^{p} \phi_i \ln \epsilon_{t-i}^2 + \sum_{j=1}^{q} \theta_j u_{t-j} + u_t, \tag{4}$$

where

$$\phi_0 = \alpha_0 + (1 - \sum_{j=1}^{q} \beta_j) \cdot E(\ln z_t^2), \tag{5}$$

$$\phi_i = \alpha_i + \beta_i, \quad 1 \leq i \leq p, \qquad \theta_j = -\beta_j, \quad 0 \leq j \leq q, \tag{6}$$

$$u_t = \ln z_t^2 - E(\ln z_t^2). \tag{7}$$

Consistent and asymptotically normal estimates of all the ARMA parameters – and hence all the log-GARCH parameters except the log-volatility intercept α_0 – is thus readily obtained via usual ARMA estimation methods subject to appropriate assumptions, see e.g. Brockwell and Davis (2006). In order to obtain an estimate of α_0 the most common solutions have been to either impose restrictive assumptions regarding the distribution of z_t (say, normality, see e.g. Psaradakis and Tzavalis (1999)), or to use an *ex post* scale-adjustment (see e.g. Bauwens and Sucarrat (2010), and Sucarrat and Escribano (2012)). What our argument below shows is that the *ex post* scale-adjustment (i.e. formula (8) below) provides a consistent estimate of $E(\ln z_t^2)$. Consequently, the final log-GARCH parameter, α_0, can also be estimated consistently.

2.3 On consistency

To obtain an understanding of the motivation behind the scale-adjustment, consider writing (1) as

$$\epsilon_t = \sigma_t^* z_t^*, \quad z_t^* \sim IID(0, \sigma_{z^*}^2),$$

where σ_t^* is a time-varying scale not necessarily equal to the standard deviation, and where z_t^* does not necessarily have unit variance. Of course, by construction $\sigma_t = \sigma_t^* \sigma_{z^*}$ and $z_t = z_t^* / \sigma_{z^*}$. Next, suppose a log-scale specification (e.g. an ARMA specification contained in (4)) is fitted to $\ln \epsilon_t^2$, with $\ln \widehat{\sigma}_t^{*2}$ denoting the fitted value of the ARMA specification such that $\widehat{\sigma}_t^* = \exp(\ln \widehat{\sigma}_t^*)$, and with the ARMA residual defined as $\widehat{u}_t = \ln \epsilon_t^2 - \ln \widehat{\sigma}_t^{*2}$. In order to obtain an estimate of the time-varying conditional standard deviation, which is needed for comparison with other volatility models, then it is natural to consider adjusting $\widehat{\sigma}_t^*$ by multiplying it with an estimate of σ_{z^*}, say, the sample standard deviation of the standardised residuals $\widehat{z}_t^*$. Although this argument is fine heuristically, it may not be apparent what underlying magnitude the adjustment in fact estimates, nor may it be straightforward to obtain the limiting properties of the adjustment under suitable conditions. In the log-GARCH model, however, the log of the scale-adjustment provides an estimate of $-E(\ln z_t^2)$. To see this consider the scale adjustment and its approximation:

$$\widehat{\sigma}^2_{z^*} = \frac{1}{T-1}\sum_{t=1}^{T}(\widehat{z}^*_t - \overline{\widehat{z}}^*_t)^2 \quad \approx \quad \frac{1}{T}\sum_{t=1}^{T}(\widehat{z}^*_t)^2 = \frac{1}{T}\sum_{t=1}^{T}\exp(\widehat{u}_t).$$

The population analogue of the final expression on the right is $E\left[\exp(u_t)\right]$. Taking the natural log of $E\left[\exp(u_t)\right]$ gives $\ln E\left[\exp(u_t)\right] = -E\left(\ln z_t^2\right)$ under the assumption that $E\left(z_t^2\right) = 1$, i.e. the identifiability assumption from (1). This suggests

$$-\ln\left[\frac{1}{T}\sum_{t=1}^{T}\exp(\widehat{u}_t)\right] \tag{8}$$

provides a consistent estimate of $E(\ln z_t^2)$ due to the continuity of the logarithm function.

The expression in square brackets in (8), i.e. $T^{-1}\sum_{t=T}\exp(\widehat{u}_t)$, is well-known as the "smearing estimate", see Duan (1983). It provides an estimate of the adjustment needed for an unbiased estimate of $E(y_t|x_t)$ when the left-hand side of the estimated model is $\ln y_t$.[4] The proof of Duan (1983), however, is for static models. In dynamic models, e.g. when the $\widehat{u}_t$'s are ARMA residuals, then a different proof strategy is needed. Complete proofs under mild assumptions that hold under all the configurations covered in this paper, however, is well beyond our scope. For simplicity and convenience, therefore, we instead formulate the set of minimal assumptions and conditions that we rely upon throughout, and only provide a proof of the key condition (**A2**) in the log-ARCH(p) case.

Formally, we rely on the following assumptions:

A1: $E\left(z_t^2\right) = 1$ and $\left|E\left(\ln z_t^2\right)\right| < \infty$.

A2: Let $\widehat{u}_t$, $t = 1, \ldots, T$, denote the ARMA-residuals resulting from estimating the ARMA representation (4). Then:

$$\frac{1}{T}\sum_{t=1}^{T}\exp(\hat{u}_t) - \frac{1}{T}\sum_{t=1}^{T}\exp(u_t) = o_P(1). \tag{9}$$

In **A1** the first moment condition is simply the identifiability condition from (1), whereas the other moment condition $\left|E\left(\ln z_t^2\right)\right| < \infty$ is required for the ARMA representation (4) to exist. For the two most commonly used densities of z_t in finance, i.e. $N(0,1)$ and t, $E\left(\ln z_t^2\right)$ is finite. Regarding **A2**, it immediately implies that (8) is a consistent estimator of $E(\ln z_t^2)$ due to the continuity of the logarithm function. As we have already noted, though, a complete proof of **A2** under all the configurations covered by this paper is beyond our scope. However, in the log-ARCH(p) case the proof is relatively straightforward.

Theorem 1. Suppose $\ln\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p}\alpha_i\ln\epsilon_{t-i}^2$ in (1)-(2), that $\ln\epsilon_t^2$ is strictly stationary and that **A1** holds. The mean-corrected AR(p) representation is then

[4] Specifically, if the estimated model is $\ln y_t = \beta'\mathbf{x}_t + u_t$ with $u_t \sim IID(0, \sigma_u^2)$, then $E(y_t|\mathbf{x}_t) = E[\exp(u_t)] \cdot \exp(\beta'\mathbf{x}_t)$.

given by $(\ln \epsilon_t^2 - E(\ln \epsilon_t^2)) = \sum_{i=1}^p \phi_i(\ln \epsilon_{t-i}^2 - E(\ln \epsilon_t^2)) + u_t$, where $\phi_i = \alpha_i$ as in (6). Define $\widetilde{Y}_t = \ln \epsilon_t^2 - T^{-1}\sum_{t=1}^T \ln \epsilon_t^2$. Let $\widehat{\phi}_1, \ldots, \widehat{\phi}_p$ denote the OLS estimates of $\phi_1, \ldots, \phi_p$ based on the $\widetilde{Y}_t$'s, let $\widehat{u}_t = \widetilde{Y}_t - \sum_{i=1}^p \widehat{\phi}_i \widetilde{Y}_{t-i}$ for $t > p$ and let $\widetilde{u}_t = 0$ for $0 < t \leq p$. If $E(z_t^4) < \infty$ and $|E[(\ln z_t^2)^2]| < \infty$, then **A2** holds.

Proof. See Appendix A. □

The Theorem states that **A2** holds when the mean-corrected AR(p) representation of a log-ARCH(p) model is estimated by OLS, which then implies that (8) is a consistent estimator of $E(\ln z_t^2)$. Next, it follows straightforwardly that all the log-ARCH(p) parameters can be estimated via the relationships (5) and (6), since $\widehat{\phi}_0 = (1 - \sum_{i=1}^p \widehat{\phi}_i) \cdot T^{-1}\sum_{t=1}^T \ln \epsilon_t^2$ provides a consistent estimate of ϕ_0 under the assumptions of the Theorem. Strict stationarity of $\ln \epsilon_t^2$ follows if the roots of the AR-polynomial are all outside the unit-circle.

2.4 On normality

Our main interest is a consistent estimator of $E(\ln z_t^2)$, so that we can use the ARMA-estimates to consistently estimate all the log-GARCH parameters via (5)-(6). To this end the limiting distribution of our estimator of $E(\ln z_t^2)$ is of minor interest. In simulations, however, the limiting distribution and an expression for the asymptotic variance can be useful in verifying simulation results.[5]

Let (8) be modified to

$$\widehat{\tau}_T = -\ln\left[\frac{1}{T}\sum_{t=1}^T \exp(\widehat{u}_t - \overline{\widehat{u}}_T)\right], \tag{10}$$

where $\overline{\widehat{u}}_T$ is the empirical mean of the ARMA-residuals. The mean-correction term $\overline{\widehat{u}}_T$ is needed, since condition **A3** (below) may not be valid without it (see e.g. the related discussion in Yu (2007), where high moment partial sum processes of residuals in ARMA models are treated). Of course, in some cases, e.g. when OLS is used to estimate the AR(p) representation of a log-ARCH(p) model, then $\overline{\widehat{u}}_T$ is zero by construction, and so (10) equals (8). The following two assumptions are needed for asymptotic normality:

A3: Let $\{\widehat{u}_t\}_{t=1}^T$ denote the ARMA-residuals resulting from estimating the ARMA representation (4). Denoting $\overline{\widehat{u}}_T$ and $\overline{u}_T$ as the averages of $\widehat{u}_t$ and u_t, respectively:

$$\sqrt{T}\left[\frac{1}{T}\sum_{t=1}^T \exp(\widehat{u}_t - \overline{\widehat{u}}_T) - \frac{1}{T}\sum_{t=1}^T \exp(u_t - \overline{u}_T)\right] = o_P(1).$$

A4: $E(z_t^4) < \infty$ and $|E[(\ln z_t^2)^2]| < \infty$.

[5] Of course, the limiting distribution is also useful for inference on $E(\ln z_t^2)$, but this is not the focus of this paper.

Condition **A3** is slightly stronger than **A2**, since **A3** implies that (10) provides a consistent estimate of $E\left(\ln z_t^2\right)$ as long as **A1** holds. The moment conditions in **A4** are needed for the asymptotic variance of (10) to be finite.

Theorem 2. Suppose (1)-(2), **A1**, **A3** and **A4** hold. Then

$$\sqrt{T}\left[\widehat{\tau}_T - E\left(\ln z_t^2\right)\right] \xrightarrow{D} N(0, \zeta^2), \tag{11}$$

where

$$\zeta^2 = Var\left(z_t^2 - \ln z^2\right). \tag{12}$$

Proof. See Appendix B. □

The key assumption for asymptotic normality to hold is **A3**, but a complete proof of it under all the configurations covered by this paper is beyond our scope. However, just as for consistency in the log-ARCH(p) case (see Theorem 1), a proof of asymptotic normality in the log-ARCH(p) case is relatively straightforward.

Theorem 3. Suppose the assumptions of Theorem 1 holds. If in addition $E\left(u_t^4\right) < \infty$, then **A3** holds.

Proof. See Appendix A. □

Assumption **A4** holds under the assumptions of Theorem 1. The additional condition $E\left(u_t^4\right) < \infty$ is in fact a very weak assumption, since it follows from $E\left(e^{|u_t|}\right) < \infty$.

An extensive set of Monte Carlo simulations have been performed, of which Table 1 only contains a small subset (more simulations are contained in Tables 2 to 6, and additional simulations are available on request). The last three columns of Table 1 confirm that the Gaussian QMLE via the ARMA representation (w/mean-correction) provides consistent estimates and empirical sample standard errors that coincide with their asymptotic counterparts. Although, as expected, a larger number of observations is needed as the persistence parameter $\phi_1 = \alpha_1 + \beta_1$ approaches 1, and when α_1 goes towards zero (i.e. a common root). Additional simulations (available on request) show similar properties for the Gaussian QMLE without mean-correction, and for the Least Squares Estimator (LSE). All simulations and computations are in *R* (R Core Team (2014)) with the `lgarch` package (Sucarrat (2014b)).

2.5 Log-GARCH-X

Additional exogenous or predetermined variables ("X") can be added linearly or non-linearly to the log-volatility specification $\ln \sigma_t^2$ without affecting the relationship between the log-GARCH coefficients and the ARMA coefficients. Specifically, let the log-GARCH-X model be given by

$$\ln \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \ln \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \ln \sigma_{t-j}^2 + g(\lambda, x_t), \tag{13}$$

where g is a linear or nonlinear function of the exogenous and/or predetermined variables x_t, and a parameter vector λ. The index t in x_t does not necessarily mean that all (or any) of its elements are contemporaneous. If $|E(\ln z_t^2)| < \infty$, then (13) admits the ARMA-X representation

$$\ln \epsilon_t^2 = \phi_0 + \sum_{i=1}^{p} \phi_i \ln \epsilon_{t-i}^2 + \sum_{j=1}^{q} \theta_j u_{t-j} + g(\lambda, x_t) + u_t, \tag{14}$$

where the ARMA coefficients are defined as before, i.e. by (5)-(6), and where u_t is the same as earlier, i.e. $u_t = \ln z_t^2 - E(\ln z_t^2)$. Rigorous proofs of consistency and asymptotic normality, which we do not provide here, would of course require precise assumptions on the behaviour of x_t, see for example Hannan and Deistler (2012, chapter 4). However, if all the ARMA-X parameters are estimated consistently, then a reasonable conjecture is that (8) provides a consistent estimate of $E(\ln z_t^2)$, and hence that all the log-GARCH parameters can be estimated consistently.

One type of conditioning variable that is of special interest in financial applications is leverage or volatility asymmetry. Table 2 provides simulation results that suggests Theorem 2 holds for a simple version of leverage, namely

$$\ln \sigma_t^2 = \alpha_0 + \alpha_1 \ln \epsilon_{t-1}^2 + \beta_1 \ln \sigma_{t-1}^2 + \lambda_1 I_{\{z_{t-1}<0\}}, \tag{15}$$

where $I_{\{z_{t-1}<0\}}$ is an indicator function equal to 1 if $z_{t-1} < 0$ and 0 otherwise. Note that $I_{\{z_{t-1}<0\}}$ is observable, since $I_{\{z_{t-1}<0\}} = I_{\{\epsilon_{t-1}<0\}}$. The simulations suggest all the parameters are estimated consistently, and the last three columns suggest the finite sample empirical standard errors of the estimate of $E(\ln z_t^2)$ correspond to their asymptotic counterparts for both the normal and the t distributions. Additional simulations are contained in Table 5, where the univariate log-GARCH-X form is used equation-by-equation to estimate a multivariate log-GARCH(1,1) model with diagonal GARCH matrix and time-varying correlations.

3 Multivariate log-GARCH

3.1 Notation and specification

The M-dimensional log-GARCH model is given by

$$\begin{aligned}
\epsilon_t &\sim ID(0, H_t), \quad t \in \mathbb{Z}, &(16)\\
D_t^2 &= \operatorname{diag}\left\{\sigma_{m,t}^2\right\}, \quad m = 1, \ldots, M, &(17)\\
z_t &= D_t^{-1}\epsilon_t, \quad \forall m : z_{m,t} \sim IID(0,1), \quad P(z_t = 0) = 0, &(18)\\
\ln \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{p} \alpha_i \ln \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \ln \sigma_{t-j}^2, \qquad p \geq q, &(19)
\end{aligned}$$

where ϵ_t, σ_t^2 and z_t are $M \times 1$ vectors, and where H_t and D_t are $M \times M$ matrices. In (19) we have that $\alpha_0 = (\alpha_{1.0}, \ldots, \alpha_{M.0})'$,

$$\alpha_i = \begin{pmatrix} \alpha_{11.i} & \cdots & \alpha_{1M.i} \\ \vdots & \ddots & \vdots \\ \alpha_{M1.i} & \cdots & \alpha_{MM.i} \end{pmatrix} \quad \text{and} \quad \beta_j = \begin{pmatrix} \beta_{11.j} & \cdots & \beta_{1M.j} \\ \vdots & \ddots & \vdots \\ \beta_{M1.j} & \cdots & \beta_{MM.j} \end{pmatrix}, \tag{20}$$

where $'$ is the transpose operator. Equation (16) means ϵ_t is independent with mean zero and a time-varying conditional covariance matrix H_t. The IID assumption in equation (18) states that each marginal series $\{z_{m,t}\}$ is $IID(0,1)$. Marginal identicality is a key characteristic of the ARCH class of models, and is needed for (8) (or (10)) to be applicable after estimation via the VARMA representation. An implication of (18) is that $z_t \sim ID(0, R_t)$, where R_t is both the conditional covariance and correlation matrix – possibly time-varying – of z_t. In other words, the vector z_t is ID but not necessarily IID, even though each marginal series $\{z_{mt}\}$ is IID. In the special case where the vector z_t is IID, then R_t is a Constant Conditional Correlation (CCC) model. Estimation of the volatilities D_t^2 does not require that the off-diagonals of H_t (i.e. the covariances) are specified explicitly. Nor need we assume that ϵ_t is distributed according to a certain density, say, the normal.

3.2 The VARMA representation

If $|E(\ln z_t^2)| < \infty$, then the M-dimensional log-GARCH(p,q) model (19) admits the VARMA(p,q) representation

$$\ln \epsilon_t^2 = \phi_0 + \sum_{i=1}^{p} \phi_i \ln \epsilon_{t-i}^2 + \sum_{j=1}^{q} \theta_j u_{t-j} + u_t, \tag{21}$$

where

$$\phi_0 = \alpha_0 + (I_M - \sum_{j=1}^{q} \beta_j) \cdot E(\ln z_t^2), \quad \phi_i = \alpha_i + \beta_i, \quad \theta_j = -\beta_j \quad \text{and} \tag{22}$$

$$u_t = \ln z_t^2 - E(\ln z_t^2). \tag{23}$$

In the special case where the vector z_t is IID, which implies a CCC model for the correlations (assuming they exist), then the vector u_t is IID as well. In this case it is well known that the multivariate Gaussian QMLE provides consistent and asymptotically normal estimates of the VARMA coefficients under suitable assumptions, see e.g. Lütkepohl (2005). Accordingly, consistent estimation and asymptotically normal inference regarding all the log-GARCH coefficients – apart from the log-volatility intercept α_0 – is available as well. In order to obtain a consistent estimate of α_0, then an estimate of the $M \times 1$ vector $E(\ln z_t^2)$ is needed. Since the process $\{u_{m,t}\}$ is marginally IID for each m, an equation-by-equation application of (8) (or of (10)) after estimation of the VARMA representation is likely to provide consistent estimates of each element in $E(\ln z_t^2)$. Tables 3 and 4 contain simulation results that support this hypothesis. The estimates of α_0 and $E(\ln z_t^2)$ are consistent, and the last two columns suggest the empirical sample standard errors coincide with their asymptotic counterparts as implied by (12).

In the case where the vector z_t is only ID, which is implied by time-varying correlations, then the vector u_t is only ID as well. This corresponds to a VARMA model with heteroscedastic error u_t. Fewer QML results are available in this case, e.g. Bardet and Wintenberger (2009). However, in the special case where the β_j matrices are diagonal, then the M-dimensional VARMA model can be estimated equation-by-equation by univariate ARMA-X methods, since – equation-by-equation – each error term $u_{m,t}$ is IID (along the lines of Francq and Zakoïan (2014)). Next, equation-by-equation application of (8) is likely to provide consistent estimates of each element in $E\left(\ln z_t^2\right)$, and hence of the log-volatility intercept α_0. Table 5 contains simulation results that supports this hypothesis when the time-varying correlations are governed by Engle's (2002) Dynamic Conditional Correlations (DCC) model. The estimates of α_0 and $E\left(\ln z_t^2\right)$ are consistent, and the last two columns suggest the empirical sample standard errors coincide with their asymptotic counterparts as implied by (12).

3.3 Multivariate log-GARCH-X

Just as in the univariate case, the multivariate log-GARCH model permits exogenous and/or predetermined conditioning variables in each of the M equations. Specifically, write the multivariate log-GARCH-X specification as

$$\ln \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \ln \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \ln \sigma_{t-j}^2 + \lambda x_t, \tag{24}$$

where x_t is an $L \times 1$ vector of predetermined or exogenous variables, and where λ is an $M \times L$ matrix. Here, for notational economy, we let the predetermined or exogenous variables x_t enter linearly, but in principle they can enter non-linearly as in the univariate case, see (14). Similarly, the index t in x_t does not necessarily mean that all (or any) of its elements are contemporaneous. The VARMA-X representation of (24) is then given by

$$\ln \epsilon_t^2 \quad = \quad \phi_0 + \sum_{i=1}^{p} \phi_i \ln \epsilon_{t-i}^2 + \sum_{j=1}^{q} \theta_j u_{t-j} + \lambda x_t + u_t,$$

with the VARMA coefficients and u_t defined as before, i.e. by (22). In other words, the relation between the VARMA coefficients and the log-GARCH coefficients are not affected by adding λx_t to (24). So VARMA-X methods can be used to estimate all the log-GARCH parameters (under suitable assumptions on x_t) except the log-volatility intercept α_0 in a first step, and then in a second step equation-by-equation application of (8) can be used to estimate each element in $E\left(\ln z_t^2\right)$ and hence the log-volatility intercept α_0. Also here it is useful to distinguish between between the CCC and time-varying correlations cases. If u_t is IID, i.e. the CCC case, then – under suitable assumptions – the multivariate Gaussian QMLE provides consistent estimates of the VARMA-X representation, see e.g. Hannan and Deistler (2012). If correlations are time-varying, and if the matrices β_j are diagonal, then each equation can be estimated separately in terms of their ARMA-X representations.

4 Application: Modelling the uncertainty of electricity prices

Short-term electricity price modelling and forecasting is of great importance for energy market participants. On the supply side, producers need forecasts of prices and the time-varying uncertainty associated with those forecasts in order to appropriately determine price and production levels. On the demand side, consumers and speculators need the same type of information to decide when and where to produce, whether to speculate and/or hedge against adverse price changes, and for risk management purposes. Daily electricity prices are characterised by autoregressive persistence, day-of-the week effects, large spikes or jumps, ARCH and non-normal conditional errors that are possibly skewed. Koopman et al. (2007), Escribano et al. (2011), and Bauwens et al. (2013) have proposed univariate and multivariate models that contain some or several of these features. However, in none of these models is the volatility specification – a non-exponential GARCH – robust to the large spikes that is a common characteristic of electricity prices (robustness is important to avoid large and persistent volatility forecast failure following spikes or "jumps"). Nor are they flexible enough to accommodate a complex and rich heteroscedasticity dynamics similar to that of the mean specification without imposing very strong parameter restrictions (e.g. non-negativity). Finally, automated model selection with a large number of variables is infeasible in practice due to computational complexity and positivity constraints. The log-GARCH-X class of models, by contrast, remedies these deficiencies. The objective of this section is to illustrate this.

4.1 Data

The data consist of the daily peak and off-peak spot electricity prices (in Euros per kw/h) from 1 January 2010 to 20 May 2014 (i.e. 1601 observations before lag-adjustments) for the Oslo region in Norway.[6] Electricity forwards for this region is traded at the Nord Pool Spot energy exchange, which is the leading European market for electrical energy. Factories, companies and other institutions with electricity consumption may want to shift part of their activity to and from peak hours for efficient cost management, since the difference between peak and off-peak prices can be very large at times, see Figure 1. As an aid in the decision-making process, forecasts of future prices and of price uncertainty (volatility) can therefore be of great usefulness. The daily peak spot price $S_{1,t}$ is computed as the average of the spot prices during peak hours, that is, $S_{1,t} = (S_{t(8am)} + \cdots + S_{t(9pm)})/14$, whereas the daily off-peak spot price $S_{2,t}$ is computed as the average of the spot prices during off-peak hours, that is, $S_{2,t} = (S_{t(0am)} + \cdots + S_{t(7am)} + S_{t(10pm)} + S_{t(11pm)})/10$. Note that $S_{t(8am)}$ should be interpreted as the electricity price from 8am to 9am, $S_{t(9am)}$ should be interpreted

[6]The source of the data is `http://www.nordpoolspot.com/`, and the sample was determined by availability: Observations prior to the sample period are not available, and the data were downloaded just after 20 May 2014.

as the electricity price from 9am to 10am, and so on. Graphs of $S_{1,t}, S_{2,t}$ and their log-returns ($r_t = \Delta \ln S_t$) are contained in Figure 1. The price and returns figures exhibit the usual characteristics of electricity prices, namely that the price variability is substantially larger than those of financial prices (say, stocks, stock indices and exchange rates), and that big jumps occur relatively frequently.

4.2 Univariate log-GARCH models

The conditional mean is specified as a two-dimensional Vector Error Correction Model (VECM) augmented with day-of-the-week dummies in both equations.[7] The residuals or mean-corrected returns from the estimated model are then used for the estimation of the log-volatility specifications. The univariate models that we fit to each of the two mean-corrected returns are

$$\text{log-GARCH(1,1)}: \quad \ln \sigma_t^2 = \alpha_0 + \alpha_1 \ln \epsilon_{t-1}^2 + \beta_1 \ln \sigma_{t-1}^2, \tag{25}$$

$$\text{log-GARCH(7,1)}: \quad \ln \sigma_t^2 = \alpha_0 + \sum_{i=1}^{7} \alpha_i \ln \epsilon_{t-i}^2 + \beta_1 \ln \sigma_{t-1}^2, \tag{26}$$

$$\text{log-GARCH(7,1)}-\text{X}: \quad \ln \sigma_t^2 = \alpha_0 + \sum_{i=1}^{7} \alpha_i \ln \epsilon_{t-i}^2 + \beta_1 \ln \sigma_{t-1}^2 + \sum_{l=1}^{6} \lambda_l x_{lt}, \tag{27}$$

$$\begin{aligned}\text{log-GARCH(7,1)}-\text{X}^*: \quad \ln \sigma_{1t}^2 =& \alpha_0 + \sum_{i=1}^{7} \alpha_{1.i} \ln \epsilon_{1,t-i}^2 + \beta_1 \ln \sigma_{1,t-1}^2 + \sum_{l=1}^{6} \lambda_l x_{lt} \\ &+ \sum_{i=1}^{7} \alpha_{2.i} \ln \epsilon_{2,t-i}^2,\end{aligned} \tag{28}$$

$$\begin{aligned}\text{log-GARCH(7,0)}-\text{X}^*: \quad \ln \sigma_{1t}^2 =& \alpha_0 + \sum_{i=1}^{7} \alpha_{1.i} \ln \epsilon_{1,t-i}^2 + \sum_{l=1}^{6} \lambda_l x_{lt} \\ &+ \sum_{i=1}^{7} \alpha_{2.i} \ln \epsilon_{2,t-i}^2,\end{aligned} \tag{29}$$

where ϵ_t is the mean-corrected return in question, and where $x_{1t}, \ldots, x_{6t}$ are six day-of-the-week dummies for Tuesday to Sunday. In the last two specifications, where we add an asterisk * to X, then $\epsilon_{2,t}$ is the mean-corrected off-peak return when $\epsilon_{1,t}$ is the mean-corrected on-peak return, and vice-versa $\epsilon_{2,t}$ is the mean-corrected on-peak return when $\epsilon_{1,t}$ is the mean-corrected off-peak return. Of course, this means the last two equations could be considered as an Equation-by-Equation-Estimation (EbEE) scheme similar to that of Francq and Zakoïan (2014) (except that we do not estimate the time-varying correlations). The last specification, i.e. log-GARCH(7,0)–X*, actually refers to a more parsimonious version than the one displayed. The parsimonious specification is obtained by automated General-to-Specific (GETS) model selection starting from (29), see Sucarrat and Escribano (2012).

Table 7 contains the estimation results of the univariate models (only a selection of

[7] The R-squared of the two equations are 0.26 and 0.17, respectively. More details are available on request.

the estimated parameters are reported for parsimony). The first striking characteristic of the results is the large ARCH(1) estimate of about 0.2 or just below for almost all the models. By contrast, daily financial returns typically exhibit an ARCH(1) estimate of about 0.05 (or lower). This means the uncertainty (i.e. volatility) of electricity returns is much more volatile in comparison. Moreover, the estimate of about 0.2 does not change much if additional variables (e.g. lags of $\ln \epsilon_t^2$ and day-of-the-week dummies) are added. By contrast, the GARCH(1) term *is* affected when additional terms are added. In the plain log-GARCH(1,1) models, for example, it is estimated to 0.64 (peak) and 0.80 (off-peak), respectively. By contrast, when additional terms are added it falls – most of the time – to about 0 or close to 0. An interesting exception to this is the log-GARCH(7,1)-X* specification of the mean-corrected peak returns, and the log-GARCH(7,1)-X specification of the mean-corrected off-peak returns. Finally, Figure 2 shows that the different specifications can produce fundamentally different volatility forecasts. In particular, the bottom graphs show that the log-GARCH(1,1) underestimates volatility on average, and that the log-GARCH(7,1)-X* models can produce fitted standard deviations that are more than twice as big. In other words, one may seriously underestimate volatility if one does not properly take lags and day-of-the-week effects into account.

4.3 Multivariate log-GARCH models

The multivariate models that we fit to the vector of mean-corrected return ϵ_t are

$$\text{m-log-GARCH(1,1)}: \quad \ln \sigma_t^2 = \alpha_0 + \alpha_1 \ln \epsilon_{t-1}^2 + \beta_1 \ln \sigma_{t-1}^2, \tag{30}$$

$$\text{m-log-GARCH(7,1)}: \quad \ln \sigma_t^2 = \alpha_0 + \sum_{i=1}^{7} \alpha_i \ln \epsilon_{t-i}^2 + \beta_1 \ln \sigma_{t-1}^2, \tag{31}$$

$$\text{m-log-GARCH(7,1)}-\text{X}^*: \quad \ln \sigma_t^2 = \alpha_0 + \sum_{i=1}^{7} \alpha_i \ln \epsilon_{t-i}^2 + \beta_1 \ln \sigma_{t-1}^2 + \lambda x_t, \tag{32}$$

where both α_i and β_1 are 2×2 matrices, x_t is a 6×1 vector containing the six day-of-the-week dummies and λ is a 2×6 matrix. Table 8 contains the estimation results of the three multivariate models (again only a selection of the estimated parameters are reported for parsimony). Just as in the univariate case the ARCH(1) estimates are considerably higher than for daily financial returns – often close to 0.2, and they do not fall when additional terms are added. The m-log-GARCH(1,1)-X* estimates might suggest that the model is not stable, since $\hat{\alpha}_{22.1} + \hat{\beta}_{22.1}$ is very close to 1. However, the roots of the lag-polynomial are in fact both outside the unit circle. Finally, also in the multivariate case is there sometimes a large difference between the fitted standard deviations. Specifically, just as in the univariate case, the plain multivariate log-GARCH(1,1) model may seriously underestimate the uncertainty (i.e. volatility) when compared with the multivariate model that also include lags and day-of-the-week periodicity in the volatility specification (i.e. m-log-GARCH(7,1)-X*). This is clearly apparent from Figure 3.

5 Conclusions

We have proposed a general and flexible framework for the estimation of and inference in univariate and multivariate Generalised log-ARCH-X (i.e. log-GARCH-X) models when the conditional density is not known. Estimation is via the (V)ARMA-X representation, which induces a bias in the log-volatility intercept made up of a log-moment expression that depends on the conditional density. We proposed an estimator of the log-moment expression, and derived its asymptotic variance under mild assumptions. Due to the structure of the problem the bias-correction procedure is likely to also hold for univariate log-GARCH-X models, and – equation-by-equation – for multivariate log-GARCH-X models. An extensive number of simulations support our conjecture. Finally, our empirical application shows that the methods are particularly useful when the volatility dynamics are complex and possibly affected by many factors.

The results in this paper suggests a vast range of new possible research questions, both empirical and theoretical. Empirically, since the methods enable a much richer and flexible approach to volatility modelling in general – both univariate and multivariate, many problems that earlier could not be handled in practice due to computational complexity are now readily implemented. Theoretically, since estimation is via the (V)ARMA representation, the vast literature on ARMA models and variants thereof serves as a source of ideas for possible extensions.

An early version of this paper (Sucarrat and Escribano (2010)) initiated the larger research agenda of which it is part. Sucarrat and Escribano (2012) relies explicitly on the results of this paper, whereas Bauwens and Sucarrat (2010) is a precursor. These papers led to the development of the *R* (R Core Team (2014)) software packages `AutoSEARCH` (Sucarrat (2012)) and `gets` (Sucarrat (2014a)) for automated General-to-Specific (Gets) modelling of log-ARCH-X models. An early critique of the log-ARCH class of models was that the log-ARCH terms in the log-volatility specification may not exist, since the errors of a regression in empirical practice can be zero. A solution to this problem, however, is proposed in Sucarrat and Escribano (2013). This solution is only available when estimation is via the (V)ARMA representation. Finally, Francq and Sucarrat (2013) propose another ARMA-based QMLE for log-GARCH models (with the centred exponential chi-squared as instrumental density) that is asymptotically more efficient when the conditional error is normal or close to normal.

References

Bardet, J.-M. and O. Wintenberger (2009). Asymptotic normality of the quasi maximum likelihood estimator for multidimensional causal processes. Unpublished working paper.

Bauwens, L., C. Hafner, and D. Pierret (2013). Multivariate Volatility Modelling of Electricity Futures. *Journal of Applied Econometrics 28*, 743–761.

Bauwens, L. and G. Sucarrat (2010). General to Specific Modelling of Exchange

Rate Volatility: A Forecast Evaluation. *International Journal of Forecasting 26*, 885–907.

Berkes, I., L. Horvath, and P. Kokoszka (2003). GARCH processes: structure and estimation. *Bernoulli 9*, 201–227.

Brockwell, P. J. and R. A. Davis (2006). *Time Series: Theory and Methods*. New York: Springer. 2nd. Edition, first published in 1991.

Brownlees, C., F. Cipollini, and G. Gallo (2012). Multiplicative Error Models. In L. Bauwens, C. Hafner, and S. Laurent (Eds.), *Handbook of Volatility Models and Their Applications*, pp. 223–247. New Jersey: Wiley.

Comte, F. and O. Lieberman (2003). Asymptotic Theory for Multivariate GARCH Processes. *Journal of Multivariate Analysis 84*, 61–84.

Duan, N. (1983). Smearing Estimate: A Nonparametric Retransformation Method. *Journal of the Americal Statistical Association 78*, pp. 605–610.

Engle, R. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflations. *Econometrica 50*, 987–1008.

Engle, R. (2002). Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business and Economic Statistics 20*, 339–350.

Engle, R. F. and T. Bollerslev (1986). Modelling the persistence of conditional variances. *Econometric Reviews 5*, 1–50.

Engle, R. F. and K. F. Kroner (1995). Multivariate simultaneous generalized ARCH. *Econometric Theory 11*, 122–150.

Escribano, A., , J. I. Peña, and P. Villaplana (2011). Modelling Electricity Prices: International Evidence. *Oxford Bulletin of Economics and Statistics 73*, 622–650.

Francq, C. and G. Sucarrat (2013). An Exponential Chi-Squared QMLE for Log-GARCH Models Via the ARMA Representation. http://mpra.ub.uni-muenchen.de/51783/.

Francq, C., O. Wintenberger, and J.-M. Zakoïan (2013). GARCH Models Without Positivity Constraints: Exponential or Log-GARCH? Forthcoming in *Journal of Econometrics*, http//dx.doi.org/10.1016/j.jeconom.2013.05.004.

Francq, C. and J.-M. Zakoïan (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli 10*, 605–637.

Francq, C. and J.-M. Zakoïan (2006). Linear-representation Based Estimation of Stochastic Volatility Models. *Scandinavian Journal of Statistics 33*, 785–806.

Francq, C. and J.-M. Zakoïan (2010). QML estimation of a class of multivariate GARCH models without moment conditions on the observed process. Unpublished working paper.

Francq, C. and J.-M. Zakoïan (2014). Estimating multivariate GARCH and stochastic correlation models equation by equation. MPRA Paper No. 54250. Online at http://mpra.ub.uni-muenchen.de/54250/.

Franses, P. H., J. Neele, and D. Van Dijk (2001). Modelling asymmetric volatility in weekly Dutch temperature data. *Environmental Modeling and Software 16*, 131–137.

Geweke, J. (1986). Modelling the Persistence of Conditional Variance: A Comment. *Econometric Reviews 5*, 57–61.

Hafner, C. and A. Preminger (2009). Asymptotic theory for a factor GARCH model. *Econometric Theory 25*, 336–363.

Hannan, E. and M. Deistler (2012). *The statistical theory of linear systems*. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM). Originally published in 1988 by Wiley, New York.

Harvey, A. C. (1976). Estimating Regression Models with Multiplicative Heteroscedasticity. *Econometrica 44*, 461–465.

Harvey, A. C. (2013). *Dynamic Models for Volatility and Heavy Tails*. New York: Cambridge University Press.

Ibragimov, R. and P. C. Phillips (2008). Regression asymptotics using martingale convergence methods. *Econometric Theory 24*(4), 888–947.

Jeantheau, T. (1998). Strong consistency of estimators for multivariate arch models. *Econometric Theory 14*, pp. 70–86.

Kawakatsu, H. (2006). Matrix exponential GARCH. *Journal of Econometrics 134*, 95–128.

Koopman, S. J., M. Ooms, and M. A. Carnero (2007). Periodic Seasonal REG-ARFIMA-GARCH Models for Daily Electricity Spot Prices. *Journal of the American Statistical Association 102*, 16–27.

Lee, S. (1997). A note on the residual empirical process in autoregressive models. *Statistics and Probability Letters 32*(4), 405–411.

Ling, S. and M. McAleer (2003). Asymptotic theory for a vector ARMA-GARCH model. *Econometric Theory 19*, 280–310.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Berlin: Springer-Verlag.

Milhøj, A. (1987). A Multiplicative Parametrization of ARCH Models. Research Report 101, University of Copenhagen: Institute of Statistics.

Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica 59*, 347–370.

Pantula, S. (1986). Modelling the Persistence of Conditional Variance: A Comment. *Econometric Reviews 5*, 71–73.

Phillips, P. C. and V. Solo (1992). Asymptotics for linear processes. *The Annals of Statistics*, 971–1001.

Psaradakis, Z. and E. Tzavalis (1999). On regression-based tests for persistence in logarithmic volatility models. *Econometric Reviews 18*, 441–448.

R Core Team (2014). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Straumann, D. and T. Mikosch (2006). Quasi-Maximum-Likelihood Estimation in Conditionally Heteroscedastic Time Series: A Stochastic Recurrence Equations Approach. *The Annals of Statistics 34*, 2449–2495.

Sucarrat, G. (2012). *AutoSEARCH: General-to-Specific (GETS) Model Selection*. R package version 1.2.

Sucarrat, G. (2014a). *gets: General-to-Specific (GETS) Model Selection*. R package version 0.2. http://cran.r-project.org/web/packages/gets/.

Sucarrat, G. (2014b). *lgarch: Simulation and estimation of log-GARCH models*.

Sucarrat, G. and Á. Escribano (2010). The Power Log-GARCH Model. Universidad Carlos III de Madrid Working Paper 10-13 in the Economic Series, June 2010. http://e-archivo.uc3m.es/bitstream/10016/8793/1/we1013.pdf.

Sucarrat, G. and Á. Escribano (2012). Automated Model Selection in Finance: General-to-Specific Modelling of the Mean and Volatility Specifications. *Oxford Bulletin of Economics and Statistics 74*, 716–735.

Sucarrat, G. and Á. Escribano (2013). Unbiased QML Estimation of Log-GARCH Models in the Presence of Zero Returns. MPRA Paper No. 50699. Online at http://mpra.ub.uni-muenchen.de/50699/.

Wintenberger, O. (2013). Continuous Invertibility and Stable QML Estimation of the EGARCH(1,1) model. *Scandinavian Journal of Statistics 40*, 846–867.

Yu, H. (2007). High moment partial sum processes of residuals in ARMA models and their applications. *Journal of Time Series Analysis 28*, 72–91.

A Proof of Theorems 1 and 3

We provide a common proof for both theorems, as they share much of the same structure.

Proof. We first note that we are here in the OLS case, which means that the residuals already have zero empirical means. Hence a mean correction is irrelevant in **A3**. In order to ease notation, we let $Y_t = \ln \epsilon_t^2$. Let $\hat{\boldsymbol{\phi}} := (\hat{\phi}_1, \dots, \hat{\phi}_p)$ be the least squares estimator of $\boldsymbol{\phi} := (\phi_1, \ldots, \phi_p)$ based on mean corrected observations $(Y_t - \bar{Y}_T)_{1\le t\le T}$ where $\bar{Y}_T = T^{-1}\sum_{t=1}^T Y_t$. Let us write $\gamma := E\, e^{u_0}$ and $\hat{\gamma} = \frac{1}{T}\sum_{t=1}^T \exp(\hat{u}_t)$. We remind the reader that (u_t) is assumed to be a zero mean IID sequence. In both Theorem 1 and 3, we are given assumption **A3**, which implies $Var\, u_0^2 = E\,[(\ln z_1^2)^2] - [E\, \ln(z_1^2)]^2 < \infty$ as well as $E\, \exp(u_1) = 1/(\exp[E\, \ln(z_1^2)]) < \infty$. When proving Theorem 3, we are also given assumption **A4**, which implies that $Var\,[\exp(u_1)] = (E\, z_1^4 - 1)/(\{\exp[E\, \ln(z_1^2)]\}^2) < \infty$, and so $E\, e^{2u_0} < \infty$.

Using this notation we see that Theorems 1 and 3 respectively follow from the following two cases which we will now show.

Case (i): If $E\, u_0^2 < \infty$ and $E\, e^{u_0} < \infty$ then $\hat{\gamma} = \gamma + o_P(1)$, i.e. $\hat{\gamma} = T^{-1}\sum_{t=1}^T \exp(u_t) + o_P(1)$.

Case (ii): If $E\, u_0^4 < \infty$ and $E\, e^{2u_0} < \infty$ then $\sqrt{T}(\hat{\gamma} - \gamma) = T^{-1/2}\sum_{t=1}^T (e^{u_t - \bar{u}_T} - \gamma) + o_P(1)$.

As a preliminary remark, we recall the standard result that $\sqrt{T}(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}') = O_P(1)$ under our assumptions (Brockwell and Davis, 2006).

We first provide some expansions that will be useful for proving both case (i) and case (ii). Let $\delta_{t,T} := \hat{u}_{t,T} - u_t$. Note that $\delta_{t,T} = -u_t$ when $t \le p$. We will for notational simplicity omit the T subscript from both $\hat{u}_{t,T}$ and $\delta_{t,T}$ in most cases. For $t \le p$ we have $e^{\hat{u}_t} = e^0 = 1$. We will see that these initial values are asymptotically insignificant and could be arbitrary. For the more interesting case $t \ge p+1$, a Taylor expansion shows that

$$e^{u_t+\delta_t} = e^{u_t} + \delta_t e^{u_t} + \delta_t^2 \int_0^1 (1-x) e^{u_t + x\delta_t} dx = e^{u_t} + \delta_t e^{u_t} + e^{u_t}\delta_t^2 \int_0^1 (1-x) e^{x\delta_t} dx. \quad (33)$$

We are therefore interested in bounding δ_t, which we will do using a simple case of the main argument in Theorem 1 of Lee (1997).

Let $\mu = E\, Y_0 = \phi_0/(1 - \sum_{i=1}^p \phi_j)$. We have that $u_t = Y_t - \mu - \sum_{i=1}^p \phi_i(Y_{t-i} - \mu)$. The definition of $\hat{u}_t$, as well as addition and subtraction shows that for $t \ge p+1$, we have that $\hat{u}_t = u_t - (\hat{\boldsymbol{\phi}} - \boldsymbol{\phi})'(Y_{t-1} - \mu, \ldots, Y_{t-p} - \mu) - T^{-1}\sum_{s=1}^T (Y_s - \mu)(1 - \sum_{i=1}^p \hat{\phi}_i)$ so that

$$\delta_{t,T} = -(\hat{\boldsymbol{\phi}} - \boldsymbol{\phi})'(Y_{t-1} - \mu, \ldots, Y_{t-p} - \mu) - T^{-1}\sum_{s=1}^T (Y_s - \mu)(1 - \sum_{i=1}^p \hat{\phi}_i). \quad (34)$$

as in Lee (1997). Lee (1997) applies the so-called Phillips-Solo device (Phillips and Solo, 1992) and concludes that

$$T^{-1/2}\sum_{s=1}^T (Y_s - \mu) = \sqrt{T}\bar{u}_T (1 - \sum_{i=1}^p \phi_i)^{-1} + \xi_T$$

with $\xi_T = o_P(1)$, see the proof of Theorem 1 in Lee (1997) immediately before his eq.(2.6). Combining this with eq. (34) implies that for $t \ge p+1$,

$$\delta_{t,T} = -(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}')(Y_{t-1} - \mu, \ldots, Y_{t-p} - \mu) - \bar{u}_T + R_T \quad (35)$$

where $R_T = o_P(T^{-1/2})$ does not depend on t. To see this, note that

$$\sqrt{T}R_T = (\sqrt{T}\bar{u}_T) - (\sqrt{T}\bar{u}_T)(1 - \sum_{i=1}^{p}\phi_i)^{-1}(1 - \sum_{i=1}^{p}\hat{\phi}_i) + \xi_T(1 - \sum_{i=1}^{p}\hat{\phi}_i)$$

Since $\sqrt{T}(\hat{\boldsymbol{\phi}} - \boldsymbol{\phi})' = O_P(1)$ we have $(1 - \sum_{i=1}^{p}\hat{\phi}_i) = (1 - \sum_{i=1}^{p}\phi_i) + o_P(1)$. Hence,

$$\begin{aligned}\sqrt{T}R_T &= (\sqrt{T}\bar{u}_T) - (\sqrt{T}\bar{u}_T)(1 - \sum_{i=1}^{p}\phi_i)^{-1}(1 - \sum_{i=1}^{p}\hat{\phi}_i) + \xi_T(1 - \sum_{i=1}^{p}\hat{\phi}_i) \\ &= (\sqrt{T}\bar{u}_T) - (\sqrt{T}\bar{u}_T)(1 + o_P(1)) + o_P(1) = o_P(1),\end{aligned}$$

where the last equality follows, since the central limit theorem implies that $\sqrt{T}\bar{u}_T = O_P(1)$. This implies that

$$\begin{aligned}M_T := \sup_{p+1\le t\le T}|\delta_{t,T}| &\le \sup_{p+1\le t\le T}|(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}')(Y_{t-1}-\mu,\dots,Y_{t-p}-\mu)| + |\bar{u}_T| + |R_T| \\ &= \sup_{p+1\le t\le T}|(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}')(Y_{t-1}-\mu,\dots,Y_{t-p}-\mu)| + |\bar{u}_T| + o_P(T^{-1/2}).\end{aligned}$$

We have that

$$\begin{aligned}T^{\alpha}\sup_{p+1\le t\le T}|(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}')(Y_{t-1}-\mu,\dots,Y_{t-p}-\mu)| &\le T^{\alpha}\sup_{1\le j\le p}|\hat{\phi}_j - \phi_j|\sup_{p+1\le t\le T}|Y_t-\mu| \\ &= \sqrt{T}\sup_{1\le j\le p}|\hat{\phi}_j - \phi_j|T^{\alpha-1/2}\sup_{p+1\le t\le T}|Y_t-\mu|.\end{aligned}$$

We now recall that $\sqrt{T}(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}') = O_P(1)$. Also, because (Y_t) is a strictly stationary linear process with exponentially decreasing coefficients, it has the same number of moments as (u_t) in the sense that $E\,u_0^{\kappa} < \infty$ implies $E\,Y_0^{\kappa} < \infty$ for any $\kappa > 0$. Suppose $0 \le \alpha < 1/2$. It is a standard result that $T^{\alpha-1/2}\sup_{p+1\le t\le T}|Y_t-\mu| = o_P(1)$ if $E\,u_0^{-1/(\alpha-1/2)} < \infty$, see e.g. Lemma 12.4 of Ibragimov and Phillips (2008). For case (i), we know $E\,u_0^2 < \infty$ which corresponds to $\alpha = 0$. For case (ii), we know $E\,u_0^4 < \infty$, corresponding to $\alpha = 1/4$. For both of these possibilities, we see that

$$T^{\alpha}\sup_{p+1\le t\le T}|(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}')(Y_{t-1}-\mu,\dots,Y_{t-p}-\mu)| = o_P(1). \tag{36}$$

Hence if $E\,u_0^2 < \infty$, the assumption we may make under case (i), we conclude that

$$M_T = \sup_{p+1\le t\le T}|(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}')(Y_{t-1}-\mu,\dots,Y_{t-p}-\mu)| + |\bar{u}_T| + o_P(T^{-1/2}) = o_P(1)$$

since $T^{-1}\sum_{t=1}^{T}u_t = o_P(1)$ by the law of large numbers.

If $E\,u_0^4 < \infty$, the assumption we may make under case (ii), we get

$$T^{1/4}M_T = T^{1/4}\sup_{p+1\le t\le T}|(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}')(Y_{t-1}-\mu,\dots,Y_{t-p}-\mu)| + T^{-1/4}|T^{1/2}\bar{u}_T| + o_P(T^{-1/4}).$$

Because $T^{1/2}\bar{u}_T = O_P(1)$ by the central limit theorem, we see that $T^{-1/4}|T^{1/2}\bar{u}_T| = o_P(1)$. By eq. (36) for $\alpha = 1/4$ we conclude that $T^{1/4}M_T = o_P(1)$.

We now show consistency, i.e. case (i). Eq. (33) shows that

$$\frac{1}{T}\sum_{t=1}^{T} e^{\hat{u}_t} = \frac{1}{T}\sum_{t=1}^{p} e^{\hat{u}_t} + \frac{1}{T}\sum_{t=q+1}^{T} e^{u_t} + \frac{1}{T}\sum_{t=q+1}^{T} \delta_t e^{u_t} + \frac{1}{T}\sum_{t=q+1}^{T} e^{u_t}\delta_t^2 \int_0^1 (1-x)e^{x\delta_t}dx, \quad (37)$$

Clearly, $\frac{1}{T}\sum_{t=1}^{p} e^{\hat{u}_t} = p/T = o_P(1)$. We have that $\int_0^1 (1-x)e^{x\delta_t}dx \leq e^{|\delta_t|}$ because for $0 \leq x \leq 1$ we have $(1-x) \leq 1$ and $e^{x\delta_t} \leq e^{|x\delta_t|} = e^{x|\delta_t|} \leq e^{|\delta_t|}$ so that $\int_0^1 (1-x)e^{x\delta_t}dx \leq \int_0^1 e^{|\delta_t|}\,\mathrm{d}x = e^{|\delta_t|}$. By $E\,e^{u_0} < \infty$, the law of large numbers implies that $\frac{1}{T}\sum_{t=p+1}^{T} e^{u_t} = E\,e^{u_0} + o_P(1)$. Hence, the triangle inequality implies that

$$|\hat{\gamma} - \gamma| \leq \frac{1}{T}\sum_{t=p+1}^{T} |\delta_t| e^{u_t} + \frac{1}{T}\sum_{t=p+1}^{T} e^{u_t}\delta_t^2 e^{|\delta_t|} + o_P(1).$$

Using $|\delta_t| \leq M_T$ we get that

$$|\hat{\gamma} - \gamma| \leq M_T \frac{1}{T}\sum_{t=p+1}^{T} e^{u_t} + M_T^2 e^{M_t} \frac{1}{T}\sum_{t=p+1}^{T} e^{u_t} + o_P(1).$$

which is $o_P(1)$ because $M_T = o_P(1)$ and $T^{-1}\sum_{t=p+1}^{T} e^{u_t} = E\,e^{u_0} + o_P(1) = O_P(1)$.

Let us now show asymptotic Normality, i.e. case (ii). From eq. (33), we see that

$$\sqrt{T}(\hat{\gamma}-\gamma) = \frac{1}{\sqrt{T}}\sum_{t=p+1}^{T} (e^{u_t}-\gamma) + \frac{1}{\sqrt{T}}\sum_{t=p+1}^{T} \delta_t e^{u_t} + \frac{1}{\sqrt{T}}\sum_{t=p+1}^{T} e^{u_t}\delta_t^2 \int_0^1 (1-x)e^{x\delta_t}dx + o_P(1)$$

The last sum is $o_P(1)$. To see this, we again use that $\int_0^1 (1-x)e^{x\delta_t}dx \leq e^{|\delta_t|}$ combined with the fact that $T^{1/4}M_T = o_P(1)$ and we see that

$$\begin{aligned}\left|\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T} e^{u_t}\delta_t^2 \int_0^1 (1-x)e^{x\delta_t}dx\right| &\leq M_T^2 e^{M_T} \frac{1}{\sqrt{T}}\sum_{t=p+1}^{T} e^{u_t} \\ &= \left(\frac{T^{1/4}}{T^{1/4}}M_T\right)^2 e^{M_T}\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T} e^{u_t} = (T^{1/4}M_T)^2 e^{M_T}\frac{1}{T}\sum_{t=p+1}^{T} e^{u_t},\end{aligned}$$

which is $o_P(1)$ because $(T^{1/4}M_T)^2 = [o_P(1)]^2 = o_P(1)$ by continuity, that $e^{M_T} = e^{o_P(1)} = e^0 + o_P(1) = 1 + o_P(1) = O_P(1)$, and by the law of large numbers $T^{-1}\sum_{t=p+1}^{T} e^{u_t} = O_P(1)$.

We have therefore shown that $\sqrt{T}(\hat{\gamma}-\gamma) = T^{-1/2}\sum_{t=p+1}^{T}(e^{u_t}-\gamma) + T^{-1/2}\sum_{t=p+1}^{T}\delta_t e^{u_t} + o_P(1) = T^{-1/2}\sum_{t=1}^{T}(e^{u_t}-\gamma) + T^{-1/2}\sum_{t=p+1}^{T}\delta_t e^{u_t} + o_P(1)$. To deal with the term including δ_t, we apply eq. (35), which implies that

$$\begin{aligned}T^{-1/2}\sum_{t=p+1}^{T}\delta_t e^{u_t} = -(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}')T^{-1/2}\sum_{t=p+1}^{T}(Y_{t-1}-\mu,\ldots,Y_{t-p}-\mu)e^{u_t} \\ - \bar{u}_T\left(T^{-1/2}\sum_{t=p+1}^{T} e^{u_t}\right) + R_T T^{-1/2}\sum_{t=p+1}^{T} e^{u_t}.\end{aligned}$$

Because Y_{t-j} and u_t are independent for $j \geq 0$, we have that $T^{-1}\sum_{t=p+1}^{T}(Y_{t-1} - \mu, \ldots, Y_{t-p}-\mu)e^{u_t} = E\,[(Y_{-1}-\mu, \ldots, Y_{-p}-\mu)e^{u_0}]+o_P(1) = (0,0,\ldots,0)E\,e^{u_0}+o_P(1) = (0,0,\ldots,0)+o_P(1)$. Hence,

$$(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}')T^{-1/2}\sum_{t=p+1}^{T}(Y_{t-1}-\mu,\ldots,Y_{t-p}-\mu)e^{u_t} = \sqrt{T}(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}')[(0,0,\ldots,0)+o_P(1)],$$

which is $o_P(1)$ because $\sqrt{T}(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}') = O_P(1)$. Recalling $R_T = o_P(T^{-1/2})$ implies that

$$R_T T^{-1/2}\sum_{t=p+1}^{T} e^{u_t} = (T^{1/2}R_T)T^{-1}\sum_{t=p+1}^{T} e^{u_t} = o_P(1)[E\,e^{u_0}+o_P(1)] = o_P(1).$$

We further have that

$$\begin{aligned}\bar{u}_T\left(T^{-1/2}\sum_{t=p+1}^{T} e^{u_t}\right) &= \sqrt{T}\bar{u}_T[E\,e^{u_0}+o_P(1)]\\ &= \sqrt{T}\bar{u}_T E\,e^{u_0} + o_P(1)\underbrace{\sqrt{T}\bar{u}_T}_{=O_P(1)} = \sqrt{T}\bar{u}_T E\,e^{u_0} + o_P(1).\end{aligned}$$

In conclusion, this shows that $\sqrt{T}(\hat{\gamma}-\gamma) = T^{-1/2}\sum_{t=1}^{T}(e^{u_t}-\gamma) - \sqrt{T}\bar{u}_T E\,e^{u_1} + o_P(1)$. We now show that $T^{-1/2}\sum_{t=1}^{T}(e^{u_t-\bar{u}_T} - E\,e^{u_0})$ fulfils exactly the same expansion. Indeed, we have that $T^{-1/2}\sum_{t=1}^{T}(e^{u_t-\bar{u}_T} - E\,e^{u_0}) = T^{-1/2}\sum_{t=1}^{T} e^{-\bar{u}_T}e^{u_t} - \sqrt{T}E\,e^{u_0} = e^{-\bar{u}_T}(T^{-1/2}\sum_{t=1}^{T} e^{u_t}) - \sqrt{T}E\,e^{u_0} = e^{-\bar{u}_T}(T^{-1/2}\sum_{t=1}^{T} e^{u_t} - E\,e^{u_0} + E\,e^{u_0}) - \sqrt{T}E\,e^{u_0} = e^{-\bar{u}_T}(T^{-1/2}\sum_{t=1}^{T}[e^{u_t} - E\,e^{u_0}]) + e^{-\bar{u}_T}\sqrt{T}E\,e^{u_0} - \sqrt{T}E\,e^{u_0} = e^{o_P(1)}(T^{-1/2}\sum_{t=1}^{T}[e^{u_t} - E\,e^{u_0}]) + [e^{-\bar{u}_T}-1]\sqrt{T}E\,e^{u_0}$. By the central limit theorem, which holds because we assume that $E\,e^{2u_0} < \infty$ we have that $T^{-1/2}\sum_{t=1}^{T}[e^{u_t} - E\,e^{u_0}] = O_P(1)$ and hence $e^{o_P(1)}T^{-1/2}\sum_{t=1}^{T}[e^{u_t} - E\,e^{u_0}] = (1+o_P(1))T^{-1/2}\sum_{t=1}^{T}[e^{u_t} - E\,e^{u_0}] = T^{-1/2}\sum_{t=1}^{T}[e^{u_t} - E\,e^{u_0}] + o_P(1)T^{-1/2}\sum_{t=1}^{T}[e^{u_t} - E\,e^{u_0}] = T^{-1/2}\sum_{t=1}^{T}[e^{u_t} - E\,e^{u_0}] + o_P(1)$. The delta method now implies that $[e^{-\bar{u}_T}-1]\sqrt{T}E\,e^{u_0} = \sqrt{T}[e^{-\bar{u}_T} - e^0]E\,e^{u_0} = -\sqrt{T}\bar{u}_T E\,e^{u_0} + o_P(1)$. The conclusion follows. $\square$

B Proof of Theorem 2

Assumption **A4** and the smoothness of the logarithm function imply that $\hat{\tau}_T$ and

$$\tilde{\tau}_T = -\ln\left[\frac{1}{T}\sum_{t=1}^{T}\exp(u_t - \bar{u}_T)\right]$$

have the same behaviour up to $o_P(T^{-1/2})$. Denoting $\tau = E\,\ln(z_1^2) = -\ln E\,e^{u_t}$, this means $\sqrt{T}(\hat{\tau}_T - \tau) = \sqrt{T}(\tilde{\tau}_T - \tau) + o_P(1)$. Slutsky's Theorem hence implies that we only need show that $\tilde{\Delta}_T = \sqrt{T}(\tilde{\tau}_T - \tau)$ is asymptotically normal. We have that

$$\tilde{\tau}_T = -\ln\frac{1}{T}\sum_{t=1}^{T} e^{u_t-\bar{u}_T} = \bar{u}_T - \ln\frac{1}{T}\sum_{t=1}^{T} e^{u_t},$$

so

$$\tilde{\Delta}_T = \sqrt{T}\bar{u}_T + \sqrt{T}\left[f\left(\frac{1}{T}\sum_{t=1}^{T} e^{u_t}\right) - f(E\, e^{u_1})\right],$$

where $f(x) = -\ln x$, with $f'(x) = -1/|x|$. By the smoothness of f, the delta method implies that

$$\begin{aligned}
\tilde{\Delta}_T &= \sqrt{T}\bar{u}_T + f'(E\, e^{u_1})\sqrt{T}\left[\frac{1}{T}\sum_{t=1}^{T} e^{u_t} - E\, e^{u_1}\right] + o_P(1) \\
&= (f'(E\, e^{u_1}), 1)\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\begin{pmatrix} e^{u_t} - E\, e^{u_1} \\ u_t \end{pmatrix} + o_P(1).
\end{aligned}$$

By the Multivariate Central Limit Theorem, we have that

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\begin{pmatrix} e^{u_t} - E\, e^{u_1} \\ u_t \end{pmatrix} \xrightarrow{d} \begin{pmatrix} X \\ Y \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} Var\, e^{u_1} & E\, u_1 e^{u_1} \\ E\, u_1 e^{u_1} & Var\, u_1 \end{pmatrix}\right)$$

where we used that $E\, u_1 = 0$ and $Cov\,(u_1, e^{u_1}) = E\, u_1 e^{u_1}$. Hence, $\tilde{\Delta}_T \xrightarrow{d} f'(E\, e^{u_1})X + Y$, which is mean zero normal with variance equal to

$$\begin{aligned}
\zeta^2 &= \left(f'(E\, e^{u_1})\right)^2\, Var\, X + Var\, Y + 2f'(E\, e^{u_1})\, Cov\,(X, Y) \\
&= \frac{Var\,[\exp(u_1)]}{[E\, \exp(u_1)]^2} + Var\,(u_1) - 2\frac{E\,[u_1 \exp(u_1)]}{E\, \exp(u_1)}.
\end{aligned}$$

Using the equalities

$$\begin{aligned}
Var\,(u_1) &= E\,[(\ln z_1^2)^2] - [E\, \ln(z_1^2)]^2 \\
Var\,[\exp(u_1)] &= \frac{1}{\{\exp[E\, \ln(z_1^2)]\}^2} \cdot (E\, z_1^4 - 1) \\
E\, \exp(u_1) &= \frac{1}{\exp[E\, \ln(z_1^2)]} \\
E\,[u_1 \exp(u_1)] &= \frac{1}{\exp[E\, \ln(z_1^2)]} \cdot \left\{E\,[(\ln z_1^2) z_1^2] - E\, \ln(z_1^2)\right\}
\end{aligned}$$

we see that

$$\zeta^2 = E\,[(\ln z_1^2)^2] - [E\,(\ln z_1^2)]^2 + (E\,(z_1^4) - 1) - 2E\,[(\ln z_1^2) z_1^2] + 2E\,(\ln z_1^2)$$

From **A4** we have that $E\,(z_1^4) < \infty$ and $E\,[(\ln z_1^2)^2] < \infty$. The Cauchy-Schwarz inequality implies that $|E\,[(\ln z_1^2) z_1^2]|^2 \leq (E\,[(\ln z_1^2)^2])(E\, z_1^4)$, so ζ^2 is finite. Finally, the expression simplifies to

$$\zeta^2 = Var\,(z_1^2 - \ln z_1^2).$$

Table 1: Finite sample properties of the Gaussian QMLE via the ARMA representation (w/mean-correction)

DGP $(\alpha_0,\alpha_1,\beta_1,\tau)$	T	$m(\hat\alpha_0)$	$se(\hat\alpha_0)$	$m(\hat\alpha_1)$	$se(\hat\alpha_1)$	$ase(\hat\alpha_1)$	$m(\hat\beta_1)$	$se(\hat\beta_1)$	$ase(\hat\beta_1)$	$m(\hat\tau)$	$se(\hat\tau)$	$ase(\hat\tau)$
$z_t \sim N(0,1)$:												
$0, 0.1, 0.8, -1.27$	1000	-0.020	0.056	0.101	0.023	0.022	0.783	0.065	0.053	-1.269	0.055	0.054
	2000	-0.009	0.034	0.100	0.015	0.016	0.794	0.040	0.038	-1.270	0.038	0.038
	5000	-0.003	0.020	0.100	0.010	0.010	0.797	0.024	0.024	-1.270	0.025	0.024
	10000	-0.003	0.015	0.100	0.007	0.007	0.797	0.017	0.017	-1.268	0.016	0.017
$0, 0.05, 0.9, -1.27$	1000	-0.041	0.149	0.052	0.018	0.016	0.865	0.116	0.040	-1.272	0.054	0.054
	2000	-0.011	0.032	0.050	0.012	0.012	0.891	0.036	0.028	-1.271	0.039	0.038
	5000	-0.004	0.015	0.050	0.008	0.007	0.896	0.020	0.018	-1.270	0.024	0.024
	10000	-0.004	0.011	0.050	0.005	0.005	0.897	0.012	0.013	-1.270	0.019	0.017
$z_t \sim t(10)$:												
$0, 0.1, 0.8, -1.39$	1000	-0.022	0.070	0.100	0.022	0.022	0.785	0.067	0.053	-1.390	0.059	0.061
	2000	-0.008	0.038	0.101	0.015	0.016	0.792	0.041	0.038	-1.392	0.044	0.043
	5000	-0.004	0.024	0.100	0.010	0.010	0.797	0.025	0.024	-1.388	0.028	0.027
	10000	0.001	0.015	0.099	0.007	0.007	0.802	0.017	0.017	-1.389	0.018	0.019
$0, 0.05, 0.9, -1.39$	1000	-0.025	0.076	0.051	0.018	0.016	0.879	0.068	0.040	-1.384	0.061	0.061
	2000	-0.011	0.032	0.050	0.012	0.012	0.891	0.033	0.028	-1.389	0.043	0.043
	5000	-0.004	0.017	0.050	0.007	0.007	0.896	0.019	0.018	-1.389	0.027	0.027
	10000	-0.002	0.012	0.050	0.005	0.005	0.899	0.012	0.013	-1.391	0.021	0.019

The estimated model is $\ln\sigma_t^2 = \alpha_0 + \alpha_1 \ln\epsilon_{t-1}^2 + \beta_1 \ln\sigma_{t-1}^2$, and estimation proceeds in three steps. First, $\mu = E(\ln\epsilon_t^2)$ is estimated with the sample mean $\hat\mu = T^{-1}\sum_{t=1}^T \ln\epsilon_t^2$. Second, an ARMA-model with ϕ_0 set to zero is fitted to the mean-corrected series $\{\ln\epsilon_t^2 - \hat\mu\}$. Third, formula (10) is used to estimate $\tau = E(\ln z_t^2)$. The ARMA estimates are then used via the relationships (5) and (6) to obtain the log-GARCH estimates. $m(x)$, sample mean of the estimate x. $se(x)$, sample standard deviation (division by R instead of $R-1$, where $R = 1000$ is the number of replications). $ase(x)$, asymptotic standard error of x (computed as $\sqrt{av(x)}/\sqrt{n}$, where $av(x)$ is the asymptotic variance of x). The expressions of $av(\hat\alpha_1)$ and $av(\hat\beta_1)$ are based on the ARMA(1,1) formulas in Brockwell and Davis (2006, pp. 259-260), whereas $av(\hat\tau) = \zeta^2$, see (12). Computations in R (R Core Team (2014)) with a developer-version of the `lgarch` package, see Sucarrat (2014b).

Table 2: Finite sample properties of the Least Squares Estimator (LSE) via the ARMA representation (without mean-correction) for a log-GARCH(1,1) with leverage

DGP $_{(\alpha_0,\alpha_1,\beta_1,\lambda_1,\tau)}$:	T	$m(\hat{\alpha}_0)$	$se(\hat{\alpha}_0)$	$m(\hat{\alpha}_1)$	$se(\hat{\alpha}_1)$	$m(\hat{\beta}_1)$	$se(\hat{\beta}_1)$	$m(\hat{\lambda}_1)$	$se(\hat{\lambda}_1)$	$m(\hat{\tau})$	$se(\hat{\tau})$	$ase(\hat{\tau})$
$z_t \sim N(0,1)$:												
$0, 0.1, 0.8, -0.01, -1.27$	1000	-0.021	0.079	0.099	0.023	0.785	0.065	-0.011	0.088	-1.271	0.054	0.054
	2000	-0.011	0.048	0.099	0.016	0.795	0.041	-0.008	0.063	-1.270	0.039	0.038
	5000	-0.004	0.028	0.100	0.010	0.797	0.024	-0.009	0.038	-1.269	0.024	0.024
	10000	-0.002	0.019	0.100	0.007	0.799	0.017	-0.010	0.026	-1.270	0.017	0.017
$0, 0.05, 0.9, -0.02, -1.27$	1000	-0.035	0.101	0.050	0.019	0.877	0.073	-0.016	0.079	-1.273	0.054	0.054
	2000	-0.013	0.045	0.050	0.012	0.891	0.039	-0.021	0.044	-1.270	0.038	0.038
	5000	-0.005	0.022	0.050	0.007	0.897	0.019	-0.020	0.028	-1.270	0.025	0.024
	10000	-0.002	0.015	0.050	0.005	0.899	0.013	-0.020	0.020	-1.270	0.017	0.017
$z_t \sim t(10)$:												
$0, 0.1, 0.8, -0.01, -1.39$	1000	-0.023	0.079	0.100	0.023	0.784	0.064	-0.010	0.094	-1.392	0.060	0.061
	2000	-0.009	0.050	0.100	0.016	0.793	0.039	-0.010	0.065	-1.391	0.043	0.043
	5000	-0.001	0.029	0.100	0.010	0.799	0.024	-0.012	0.038	-1.390	0.027	0.027
	10000	-0.003	0.022	0.100	0.007	0.798	0.017	-0.010	0.027	-1.391	0.019	0.019
$0, 0.05, 0.9, -0.02, -1.39$	1000	-0.038	0.119	0.050	0.018	0.874	0.090	-0.027	0.078	-1.392	0.061	0.061
	2000	-0.016	0.051	0.050	0.013	0.889	0.040	-0.022	0.049	-1.390	0.045	0.043
	5000	-0.004	0.024	0.050	0.008	0.897	0.019	-0.021	0.030	-1.390	0.027	0.027
	10000	-0.002	0.016	0.050	0.005	0.899	0.013	-0.021	0.021	-1.391	0.019	0.019

The estimated model is $\ln\sigma_t^2 = \alpha_0 + \alpha_1 \ln\epsilon_{t-1}^2 + \beta_1 \ln\sigma_{t-1}^2 + \lambda_1 I_{\{z_{t-1}<0\}}$, and estimation proceeds in two steps. First, the ARMA-representation $\ln\epsilon_t^2 = \phi_0 + \phi_1 \ln\epsilon_{t-1}^2 + \theta_1 u_{t-1} + \lambda I_{\{z_{t-1}<0\}} + u_t$ is fitted by the LSE. Second, formula (10) is used to estimate $\tau = E(\ln z_t^2)$. Next, the ARMA estimates are used via the relationships (5) and (6) to obtain the log-GARCH estimates. $m(x)$, sample mean of the estimate x. $se(x)$, sample standard deviation (division by R instead of $R-1$, where $R = 1000$ is the number of replications). $ase(\hat{\tau})$, asymptotic standard error of $\hat{\tau}$, see Table 1. Computations in R (R Core Team (2014)) with the `lgarch` package version 0.2, see Sucarrat (2014b).

Table 3: Finite sample properties of multivariate Gaussian QML via the VARMA representation of a 2-dimensional CCC-log-GARCH(1,1): DGP no. 1

DGP1	T	$m(\hat{\alpha}_{1.0})$	$se(\hat{\alpha}_{1.0})$	$m(\hat{\alpha}_{11.1})$	$se(\hat{\alpha}_{11.1})$	$m(\hat{\alpha}_{12.1})$	$se(\hat{\alpha}_{12.1})$	$m(\hat{\beta}_{11.1})$	$se(\hat{\beta}_{11.1})$	$m(\hat{\beta}_{12.1})$	$se(\hat{\beta}_{12.1})$	$m(\hat{\tau}_1)$	$se(\hat{\tau}_1)$	$ase(\hat{\tau}_1)$
Eq. 1:	1000	-0.018	0.088	0.098	0.024	0.001	0.024	0.784	0.076	0.003	0.074	-1.272	0.057	0.054
	2000	-0.011	0.052	0.098	0.016	0.000	0.017	0.794	0.045	0.000	0.044	-1.272	0.038	0.038
	5000	-0.003	0.028	0.100	0.010	0.000	0.010	0.797	0.025	0.001	0.025	-1.271	0.025	0.024
	10000	-0.003	0.020	0.100	0.007	0.000	0.007	0.799	0.017	0.000	0.018	-1.271	0.017	0.017
	T	$m(\hat{\alpha}_{2.0})$	$se(\hat{\alpha}_{2.0})$	$m(\hat{\alpha}_{21.1})$	$se(\hat{\alpha}_{21.1})$	$m(\hat{\alpha}_{22.1})$	$se(\hat{\alpha}_{22.1})$	$m(\hat{\beta}_{21.1})$	$se(\hat{\beta}_{21.1})$	$m(\hat{\beta}_{22.1})$	$se(\hat{\beta}_{22.1})$	$m(\hat{\tau}_2)$	$se(\hat{\tau}_2)$	$ase(\hat{\tau}_2)$
Eq. 2:	1000	-0.020	0.085	0.001	0.024	0.097	0.025	-0.003	0.073	0.791	0.068	-1.268	0.055	0.054
	2000	-0.010	0.049	0.000	0.016	0.099	0.016	-0.001	0.042	0.793	0.043	-1.270	0.036	0.038
	5000	-0.006	0.028	0.000	0.010	0.100	0.010	-0.001	0.025	0.797	0.025	-1.272	0.024	0.024
	10000	-0.002	0.020	0.000	0.007	0.100	0.007	0.000	0.017	0.799	0.017	-1.270	0.017	0.017

The estimated model is $\ln\sigma_t^2 = \alpha_0 + \alpha_1 \ln\epsilon_{t-1}^2 + \beta_1 \ln\sigma_{t-1}^2$, where $\alpha_0 = (0,0)'$, $\alpha_1 = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix}$, $\beta_1 = \begin{pmatrix} 0.8 & 0 \\ 0 & 0.8 \end{pmatrix}$ and $Corr(z_{1t}, z_{2t}) = 0.3$. Estimation proceeds in three steps. (Note: The correlation $Corr(z_{1t}, z_{2t})$ is not estimated.) First, the VARMA representation is estimated with the multivariate Gaussian QMLE. Second, the VARMA residuals are used equation-by-equation to estimate $\tau_1 = E(\ln z_{1t}^2)$ and $\tau_2 = E(\ln z_{2t}^2)$, respectively, with formula (10). Finally, the VARMA estimates and $\widehat{\tau}_1$ and $\widehat{\tau}_2$ are combined using the relationships in (22) to obtain the log-GARCH estimates. $m(x)$, sample mean of the estimate x. $se(x)$, sample standard deviation (division by R instead of $R-1$, where $R = 1000$ is the number of replications). $ase(x)$, asymptotic standard error of x (computed as $\sqrt{av(x)}/\sqrt{T}$, where $av(\hat{\tau}_1) = av(\hat{\tau}_2) = \zeta^2$, see (12)). Computations in R (R Core Team (2014)) with the `lgarch` package version 0.3, see Sucarrat (2014b).

Table 4: Finite sample properties of multivariate Gaussian QML via the VARMA representation of a 2-dimensional CCC-log-GARCH(1,1): DGP no. 2 and 3

DGP2	T	$m(\hat{\alpha}_{1.0})$	$se(\hat{\alpha}_{1.0})$	$m(\hat{\alpha}_{11.1})$	$se(\hat{\alpha}_{11.1})$	$m(\hat{\alpha}_{12.1})$	$se(\hat{\alpha}_{12.1})$	$m(\hat{\beta}_{11.1})$	$se(\hat{\beta}_{11.1})$	$m(\hat{\beta}_{12.1})$	$se(\hat{\beta}_{12.1})$	$m(\hat{\tau}_1)$	$se(\hat{\tau}_1)$	$ase(\hat{\tau}_1)$
Eq. 1:	1000	-0.020	0.135	0.095	0.029	0.049	0.029	0.683	0.170	0.125	0.269	-1.277	0.056	0.054
	2000	-0.006	0.071	0.100	0.019	0.050	0.021	0.678	0.107	0.131	0.168	-1.270	0.038	0.038
	5000	-0.005	0.043	0.100	0.012	0.050	0.012	0.695	0.058	0.106	0.083	-1.270	0.024	0.024
	10000	-0.002	0.027	0.100	0.008	0.050	0.009	0.698	0.038	0.102	0.056	-1.271	0.016	0.017
	T	$m(\hat{\alpha}_{2.0})$	$se(\hat{\alpha}_{2.0})$	$m(\hat{\alpha}_{21.1})$	$se(\hat{\alpha}_{21.1})$	$m(\hat{\alpha}_{22.1})$	$se(\hat{\alpha}_{22.1})$	$m(\hat{\beta}_{21.1})$	$se(\hat{\beta}_{21.1})$	$m(\hat{\beta}_{22.1})$	$se(\hat{\beta}_{22.1})$	$m(\hat{\tau}_2)$	$se(\hat{\tau}_2)$	$ase(\hat{\tau}_2)$
Eq. 2:	1000	-0.021	0.164	-0.001	0.030	0.097	0.031	0.127	0.190	0.539	0.270	-1.269	0.056	0.054
	2000	-0.007	0.092	0.000	0.020	0.099	0.021	0.112	0.119	0.576	0.170	-1.269	0.038	0.038
	5000	-0.001	0.053	-0.001	0.013	0.099	0.013	0.106	0.067	0.592	0.096	-1.270	0.024	0.024
	10000	0.000	0.035	0.000	0.009	0.100	0.009	0.104	0.044	0.594	0.063	-1.270	0.017	0.017

DGP3	T	$m(\hat{\alpha}_{1.0})$	$se(\hat{\alpha}_{1.0})$	$m(\hat{\alpha}_{11.1})$	$se(\hat{\alpha}_{11.1})$	$m(\hat{\alpha}_{12.1})$	$se(\hat{\alpha}_{12.1})$	$m(\hat{\beta}_{11.1})$	$se(\hat{\beta}_{11.1})$	$m(\hat{\beta}_{12.1})$	$se(\hat{\beta}_{12.1})$	$m(\hat{\tau}_1)$	$se(\hat{\tau}_1)$	$ase(\hat{\tau}_1)$
Eq. 1:	1000	-0.010	0.220	0.096	0.029	0.049	0.030	0.653	0.253	0.151	0.263	-1.274	0.054	0.054
	2000	-0.012	0.136	0.099	0.020	0.050	0.020	0.648	0.213	0.151	0.216	-1.271	0.041	0.038
	5000	-0.004	0.066	0.100	0.012	0.050	0.012	0.683	0.118	0.117	0.119	-1.271	0.025	0.024
	10000	-0.002	0.041	0.100	0.008	0.050	0.008	0.696	0.072	0.104	0.073	-1.271	0.018	0.017
	T	$m(\hat{\alpha}_{2.0})$	$se(\hat{\alpha}_{2.0})$	$m(\hat{\alpha}_{21.1})$	$se(\hat{\alpha}_{21.1})$	$m(\hat{\alpha}_{22.1})$	$se(\hat{\alpha}_{22.1})$	$m(\hat{\beta}_{21.1})$	$se(\hat{\beta}_{21.1})$	$m(\hat{\beta}_{22.1})$	$se(\hat{\beta}_{22.1})$	$m(\hat{\tau}_2)$	$se(\hat{\tau}_2)$	$ase(\hat{\tau}_2)$
Eq. 2:	1000	-0.031	0.246	0.049	0.027	0.097	0.030	0.168	0.264	0.629	0.269	-1.269	0.054	0.054
	2000	-0.011	0.133	0.050	0.020	0.099	0.020	0.145	0.214	0.653	0.209	-1.273	0.037	0.038
	5000	-0.007	0.068	0.050	0.012	0.100	0.012	0.113	0.122	0.685	0.123	-1.269	0.025	0.024
	10000	-0.003	0.039	0.050	0.008	0.100	0.008	0.106	0.069	0.694	0.070	-1.271	0.017	0.017

Notes: See Table 3. DPG2: $\alpha_1 = c(0,0)'$, $\alpha_1 = \begin{pmatrix} 0.10 & 0 \\ 0.05 & 0.10 \end{pmatrix}$, $\beta_1 = \begin{pmatrix} 0.7 & 0.1 \\ 0.1 & 0.6 \end{pmatrix}$ and $Corr(z_{1t}, z_{2t}) = 0.2$. DPG3: $\alpha_1 = c(0,0)'$, $\alpha_1 = \begin{pmatrix} 0.10 & 0.05 \\ 0.05 & 0.10 \end{pmatrix}$, $\beta_1 = \begin{pmatrix} 0.7 & 0.1 \\ 0.1 & 0.7 \end{pmatrix}$ and $Corr(z_{1t}, z_{2t}) = 0.1$. Computations in R (R Core Team (2014)) with the `lgarch` package version 0.3, see Sucarrat (2014b).

Table 5: Finite sample properties of equation-by-equation Gaussian QML (without mean-correction) of a 2-dimensional log-GARCH(1,1) w/diagonal matrix β_1 when the correlations follow the DCC of Engle (2002)

	T	$m(\widehat{\alpha}_{10})$	$m(\widehat{\alpha}_{20})$	$m(\widehat{\alpha}_{11})$	$m(\widehat{\alpha}_{21})$	$m(\widehat{\alpha}_{12})$	$m(\widehat{\alpha}_{22})$	$m(\widehat{\beta}_{11})$	$m(\widehat{\beta}_{22})$	$m(\widehat{\tau}_1)$	$se(\widehat{\tau}_1)$	$m(\widehat{\tau}_2)$	$se(\widehat{\tau}_2)$	$ase(\widehat{\tau})$
DGP1:	1000	-0.065	-0.229	0.046	0.101	0.101	0.048	0.902	0.680	-1.270	0.056	-1.270	0.054	0.054
	2000	-0.032	-0.109	0.048	0.101	0.101	0.049	0.901	0.690	-1.271	0.038	-1.270	0.039	0.038
	5000	-0.013	-0.042	0.049	0.100	0.100	0.049	0.900	0.697	-1.271	0.024	-1.270	0.023	0.024
	10000	-0.005	-0.023	0.049	0.100	0.100	0.050	0.900	0.698	-1.271	0.017	-1.270	0.017	0.017
DGP2:	1000	-0.029	-0.026	0.098	0.053	0.053	0.097	0.791	0.792	-1.270	0.055	-1.268	0.053	0.054
	2000	-0.019	-0.013	0.099	0.051	0.051	0.099	0.794	0.797	-1.271	0.038	-1.272	0.039	0.038
	5000	-0.005	-0.004	0.100	0.051	0.050	0.099	0.799	0.799	-1.270	0.024	-1.270	0.024	0.024
	10000	-0.003	-0.002	0.100	0.050	0.050	0.100	0.799	0.799	-1.269	0.017	-1.271	0.017	0.017

The estimated model is $\ln\sigma_t^2 = \alpha_0 + \alpha_1 \ln\epsilon_{t-1}^2 + \beta_1 \ln\sigma_{t-1}^2$, where $\alpha_0 = (\alpha_{10}, \alpha_{20})'$, $\alpha_1 = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}$ and $\beta_1 = \mathrm{diag}(\beta_{11}, \beta_{22})$. The standardised errors $(z_{1t}, z_{2t})'$ are governed by an Engle (2002) DCC given by $(z_{1t}, z_{2t})' \sim N(0, \Sigma_t)$, $\Sigma_t = \begin{pmatrix} 1 & \rho_t \\ \rho_t & 1 \end{pmatrix}$, $\rho_t = q_{12,t}/\sqrt{q_{1,t}q_{2,t}}$, $q_{12,t} = \overline{\rho} + a(z_{1,t-1}z_{2,t-1} - \overline{\rho}) + b(q_{12,t} - \overline{\rho})$, $q_{1,t} = 1 + a(z_{1,t-1}^2 - 1) + b(q_{1,t} - 1)$, $q_{2,t} = 1 + a(z_{2,t-1}^2 - 1) + b(q_{2,t} - 1)$ with $a = 0.05$ and $b = 0.9$. Estimation proceeds in three steps (the Engle (2002) DCC is not estimated). First, a univariate ARMA-X specification is fitted to each of the two equations with the Gaussian QMLE. Second, the ARMA-X residuals $\widehat{u}_{1t}$ and $\widehat{u}_{2t}$, respectively, are used equation-by-equation to estimate τ_1 and τ_2, respectively, with formula (10). Finally, the ARMA-X estimates and $\widehat{\tau}_1$ and $\widehat{\tau}_2$ are combined using the relationships in (22) to obtain the log-GARCH estimates. $m(x)$, sample mean of the estimate x. $se(x)$, sample standard deviation (division by R instead of $R-1$, where $R = 1000$ is the number of replications). $ase(x)$, asymptotic standard error of x (computed as $\sqrt{av(x)}/\sqrt{T}$, where $av(\hat{\tau}_1) = av(\hat{\tau}_2) = \zeta^2$, see (12)). In DGP no. 1: $\alpha_1 = c(0,0)'$, $\alpha_1 = \begin{pmatrix} 0.05 & 0.10 \\ 0.10 & 0.05 \end{pmatrix}$, $\beta_1 = \mathrm{diag}(0.90, 0.70)$ and $\overline{\rho} = -0.2$. In DGP no. 2: $\alpha_1 = c(0,0)'$, $\alpha_1 = \begin{pmatrix} 0.10 & 0.05 \\ 0.05 & 0.10 \end{pmatrix}$, $\beta_1 = \mathrm{diag}(0.80, 0.80)$ and $\overline{\rho} = 0.4$. Computations in R (R Core Team (2014)) with the `lgarch` package version 0.2, see Sucarrat (2014b).

Table 6: Finite sample properties of multivariate Gaussian QML via the VARMA-X representation of a multivariate CCC-log-GARCH-X

DGP1	T	$m(\hat{\alpha}_{1.0})$	$se(\hat{\alpha}_{1.0})$	$m(\hat{\alpha}_{11.1})$	$m(\hat{\alpha}_{12.1})$	$m(\hat{\beta}_{11.1})$	$m(\hat{\beta}_{12.1})$	$m(\hat{\lambda}_{11})$	$se(\hat{\lambda}_{11})$	$m(\hat{\tau}_1)$	$se(\hat{\tau}_1)$	$ase(\hat{\tau}_1)$
Eq. 1:	1000	-0.033	0.140	0.093	0.048	0.647	0.157	0.094	0.046	-1.272	0.054	0.054
	2000	-0.020	0.097	0.098	0.050	0.668	0.130	0.097	0.033	-1.272	0.040	0.038
	5000	-0.008	0.044	0.099	0.050	0.688	0.112	0.099	0.018	-1.270	0.023	0.024
	10000	-0.003	0.029	0.099	0.050	0.695	0.105	0.099	0.013	-1.271	0.017	0.017
	T	$m(\hat{\alpha}_{2.0})$	$se(\hat{\alpha}_{2.0})$	$m(\hat{\alpha}_{21.1})$	$m(\hat{\alpha}_{22.1})$	$m(\hat{\beta}_{21.1})$	$m(\hat{\beta}_{22.1})$	$m(\hat{\lambda}_{21})$	$se(\hat{\lambda}_{21})$	$m(\hat{\tau}_2)$	$se(\hat{\tau}_2)$	$ase(\hat{\tau}_2)$
Eq. 2:	1000	-0.006	0.128	0.049	0.095	0.138	0.668	0.204	0.045	-1.274	0.056	0.054
	2000	-0.003	0.086	0.050	0.097	0.125	0.679	0.203	0.032	-1.271	0.039	0.038
	5000	0.001	0.044	0.050	0.099	0.110	0.692	0.201	0.018	-1.269	0.024	0.024
	10000	0.000	0.028	0.050	0.099	0.105	0.695	0.201	0.012	-1.271	0.018	0.017

The estimated model is $\ln\sigma_t^2 = \alpha_0 + \alpha_1 \ln\epsilon_{t-1}^2 + \beta_1 \ln\sigma_{t-1}^2 + \lambda_1 x_t$, where $\alpha_0 = (0,0)'$, $\alpha_1 = \begin{pmatrix} 0.1 & 0.05 \\ 0.05 & 0.1 \end{pmatrix}$, $\beta_1 = \begin{pmatrix} 0.7 & 0.1 \\ 0.1 & 0.7 \end{pmatrix}$, $\lambda_1 = (0.1, 0.2)'$ and $Corr(z_{1t}, z_{2t}) = 0.4$. The variable x_t is governed by an the exogenous AR(1) process: $x_t = 0.5x_{t-1} + u_{xt}$ with $u_{xt} \overset{IID}{\sim} N(0,1)$. Estimation proceeds in three steps (the correlation $Corr(z_{1t}, z_{2t})$ is not estimated). First, the VARMA-X representation is estimated with the multivariate Gaussian QMLE. Second, the VARMA residuals are used equation-by-equation to estimate τ_1 and τ_2, respectively, with formula (10). Finally, the VARMA estimates and $\widehat{\tau}_1$ and $\widehat{\tau}_2$ are combined using (22) to obtain the log-GARCH estimates. $m(x)$, sample mean of the estimate x. $se(x)$, sample standard deviation (division by R instead of $R-1$, where $R = 1000$ is the number of replications). $ase(x)$, asymptotic standard error of x (computed as $\sqrt{av(x)}/\sqrt{T}$, where $av(\hat{\tau}_1) = av(\hat{\tau}_2) = \zeta^2$, see (12)). Computations in R (R Core Team (2014)) with the `lgarch` package, see Sucarrat (2014b).

Table 7: Estimation results of the univariate models (25)-(29) (only selected parameters are reported)

	Model	$\hat{\alpha}_0$ (*s.e.*)	$\hat{\alpha}_1$ (*s.e.*)	$\hat{\beta}_1$ (*s.e.*)	$\hat{\tau}$ (*s.e.*)	*LogL*	k
Peak:	log-GARCH(1,1)	−0.434	0.202 (0.03)	0.639 (0.06)	−1.95 (0.14)	1890.3	3
	log-GARCH(7,1)	−0.976	0.232 (0.03)	−0.039 (0.20)	−2.01 (0.18)	1841.9	9
	log-GARCH(7,1)-X	−0.127	0.228 (0.03)	0.014 (0.29)	−1.94 (0.15)	1896.4	15
	log-GARCH(7,1)-X*	0.850	0.200 (0.03)	0.798 (0.03)	−1.83 (0.12)	1989.0	22
	log-GARCH(7,0)-X*	−0.071	0.209 (0.03)	–	−1.87 (0.12)	1955.5	13
Off-peak:	log-GARCH(1,1)	−0.070	0.137 (0.02)	0.792 (0.03)	−2.03 (0.10)	1676.0	3
	log-GARCH(7,1)	−0.548	0.202 (0.03)	−0.103 (0.14)	−2.05 (0.10)	1665.9	9
	log-GARCH(7,1)-X	−0.656	0.199 (0.03)	0.801 (0.03)	−1.87 (0.08)	1807.4	15
	log-GARCH(7,1)-X*	0.129	0.163 (0.03)	−0.041 (0.33)	−1.81 (0.07)	1850.0	22
	log-GARCH(7,0)-X*	−0.047	0.179 (0.03)	–	−1.86 (0.08)	1812.6	13

$\hat{\tau}$, estimate of $E\left(\ln z_t^2\right)$. *s.e.*, standard error of estimate. *LogL*, Gaussian log-likelihood computed as $\sum_{t=1}^{T} \ln f_\epsilon(\epsilon_t; \hat{\sigma}_t)$, where $f_\epsilon(\epsilon_t; \hat{\sigma}_t)$ is the univariate normal density, ϵ_t is the mean-corrected return and $\hat{\sigma}_t$ is the fitted standard deviation ($T = 1586$ is the number of observations). k, the total number of log-GARCH parameters (τ not included). Estimation of the ARMA representation is with the LSE without mean-correction. Computations in R (R Core Team (2014)) with the `lgarch` and `AutoSEARCH` packages, see Sucarrat (2014b, 2012).

Table 8: Estimation results of the multivariate models (30)-(32) (only selected parameters are reported)

	Equation	$\hat{\alpha}_{m0}$ $(s.e.)$	$\hat{\alpha}_{mm.1}$ $(s.e.)$	$\hat{\beta}_{mm.1}$ $(s.e.)$	$\hat{\tau}_m$ $(s.e.)$	$LogL$	k
m-log-GARCH(1,1)	Peak:	−0.344	0.165 (0.03)	0.642 (0.07)	−1.85 (0.10)	4095.7	10
	Off-peak:	−0.220	0.113 (0.02)	0.887 (0.02)	−1.96 (0.10)		
m-log-GARCH(7,1)	Peak eq:	−0.053	0.193 (0.03)	0.807 (0.03)	−1.874 (0.13)	4017.1	34
	Off-peak eq:	0.128	0.170 (0.03)	0.010 (0.165)	−2.02 (0.11)		
m-log-GARCH(7,1)-X*	Peak eq:	0.143	0.206 (0.02)	0.163 (0.02)	−1.85 (0.12)	4316.4	46
	Off-peak eq:	−0.296	0.160 (0.02)	0.840 (0.02)	−1.76 (0.07)		

$\hat{\tau}_m$, estimate of $E\left(\ln z_{m,t}^2\right)$. $s.e.$, standard error of estimate. $LogL$, Gaussian log-likelihood computed as $\sum_{t=1}^{T} \ln f_\epsilon(\epsilon_t; \hat{\sigma}_t, \hat{R})$, where $f_\epsilon(\epsilon_t; \hat{\sigma}_t, \hat{R})$ is the multivariate normal density, ϵ_t is the vector of mean-corrected returns, $\hat{\sigma}_t$ is the vector of fitted standard deviations and $\hat{R}$ is the sample correlation matrix of $\hat{z}_t$ ($T = 1586$ is the number of observations). k, the total number of log-GARCH parameters from the multivariate model (τ_1 and τ_2 are not included). Estimation of the VARMA representation is with the multivariate Gaussian QMLE without mean-correction. Computations in R (R Core Team (2014)) with the `lgarch` package, see Sucarrat (2014b).

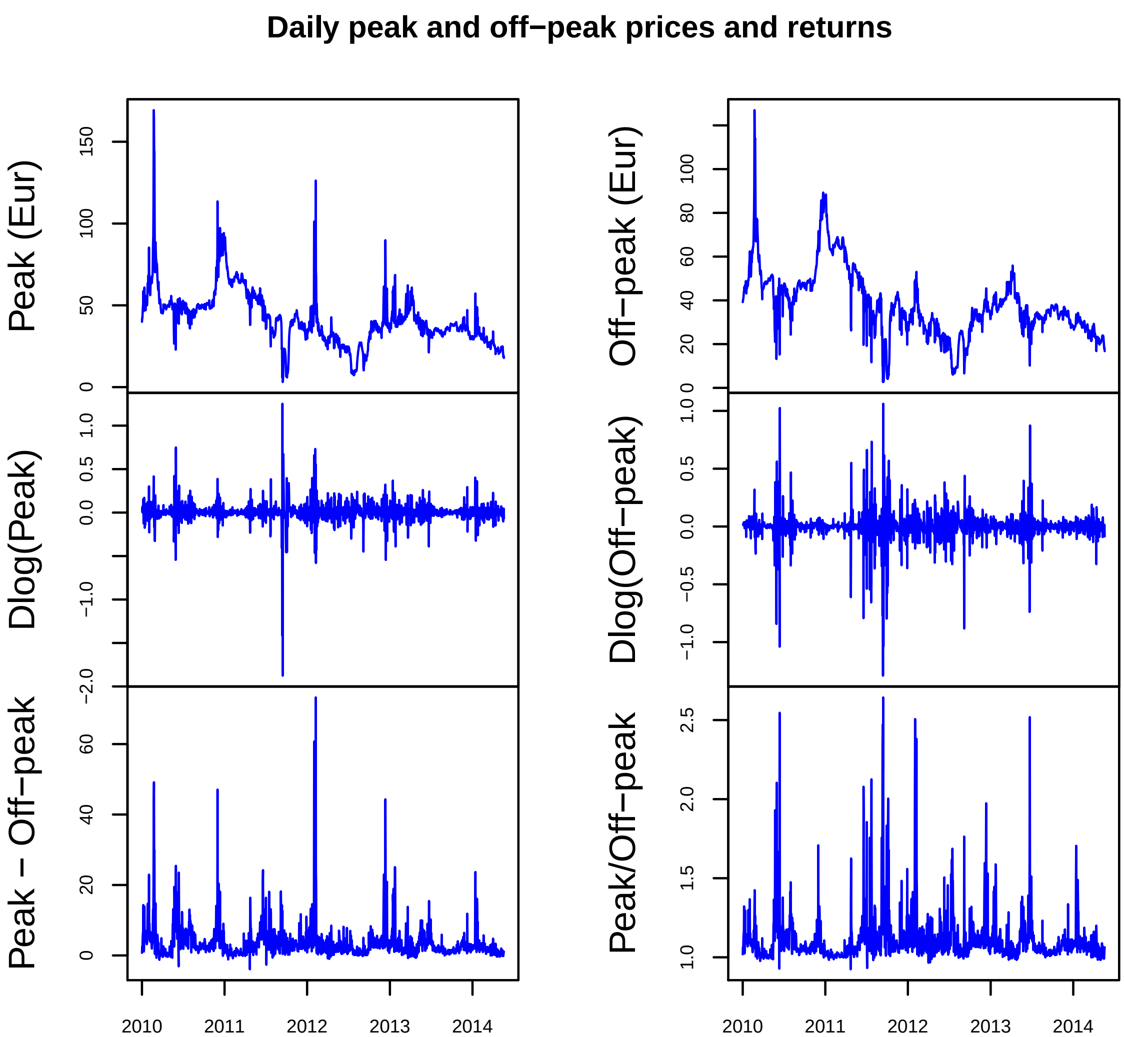

Figure 1: Daily peak and off-peak spot electricity prices (and their nominal and relative differences) in Euros per Mw/h, and log-returns for the Oslo area in Norway, 1 January 2010 - 20 May 2014 (1601 observations before lag-adjustments)

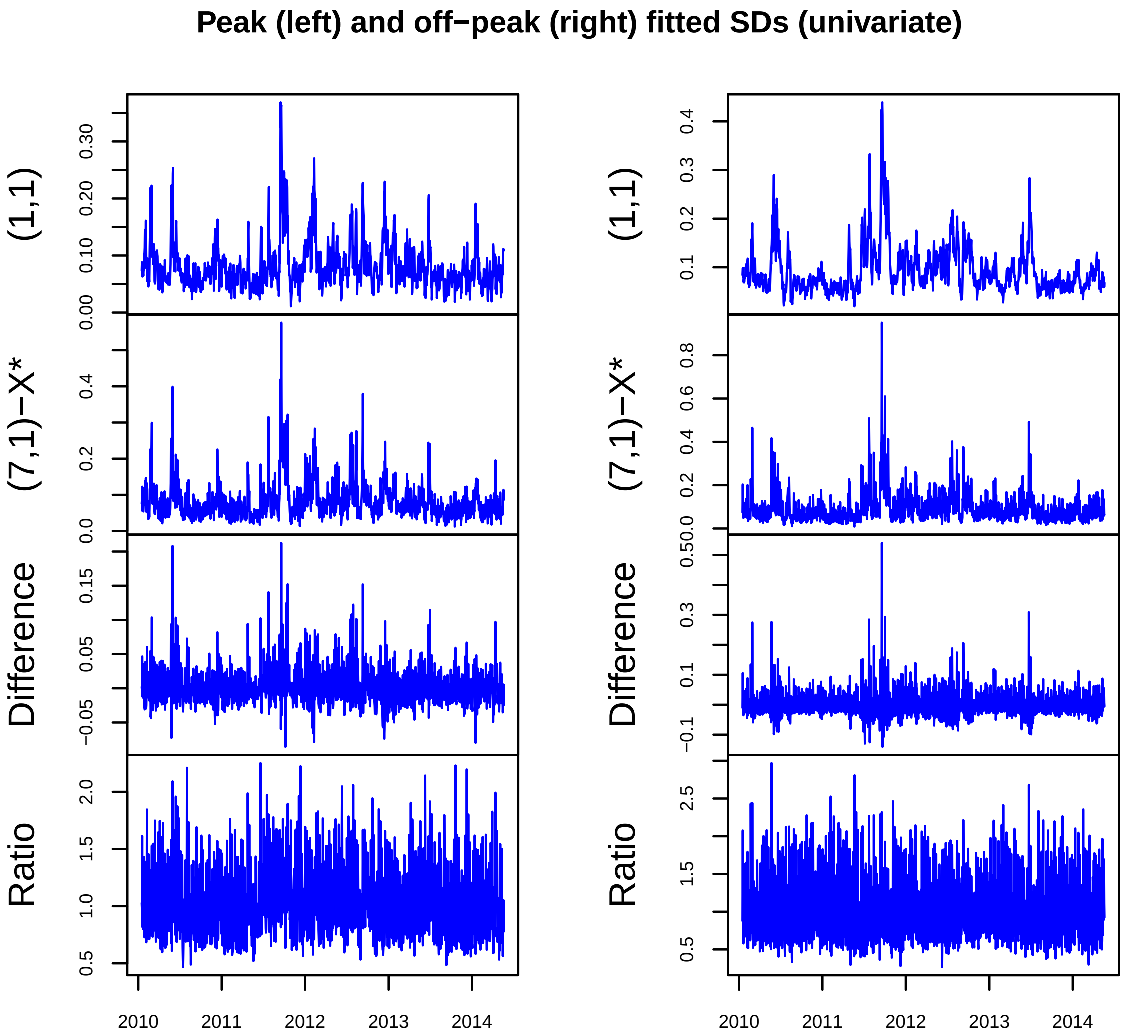

Figure 2: Fitted standard deviations (SDs) of the univariate log-GARCH(1,1) and log-GARCH(7,1)-X* models, and the nominal and relative differences between the SDs (computed as log-GARCH(1,1) minus log-GARCH(7,1)-X* and log-GARCH(1,1) over log-GARCH(7,1)-X*, respectively)

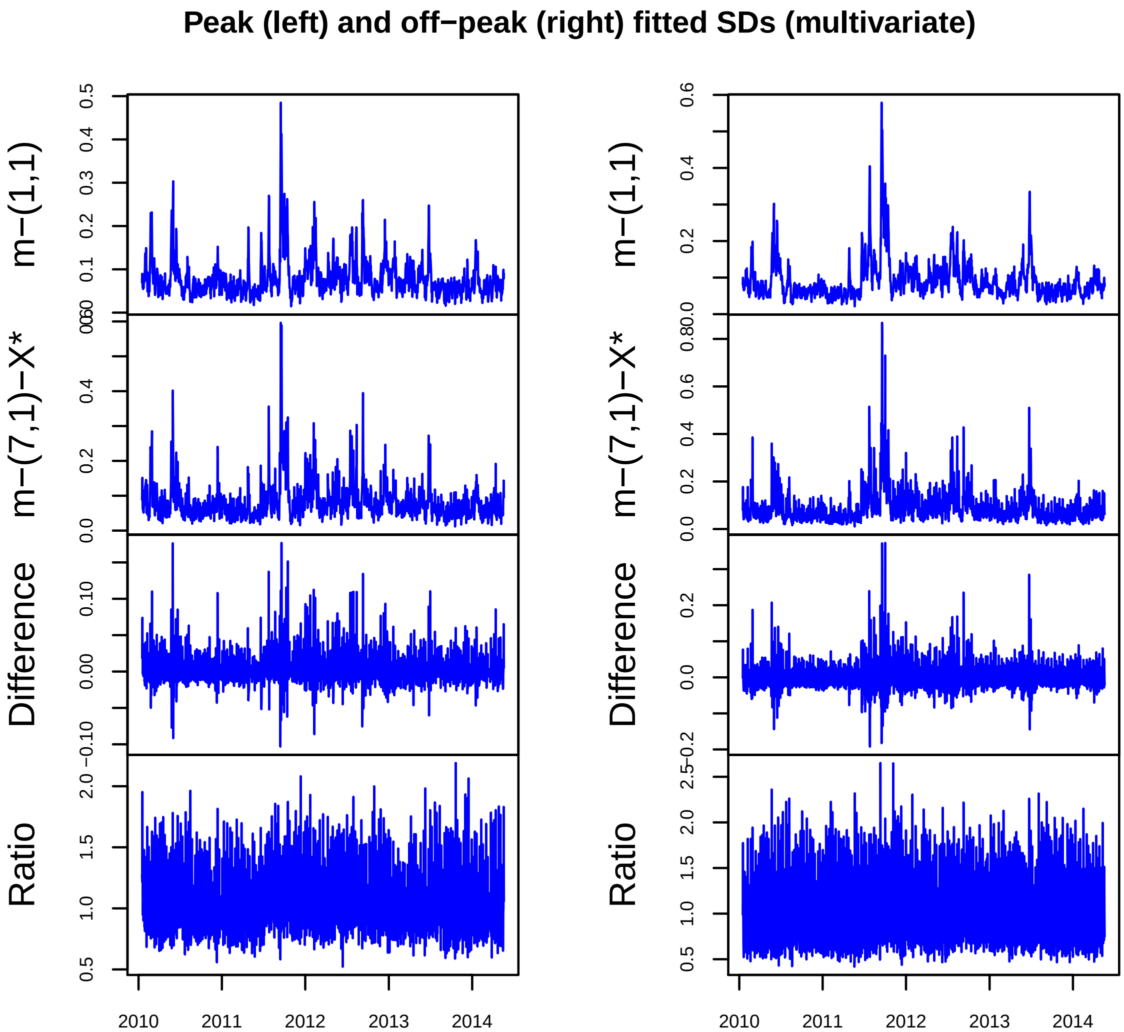

Figure 3: Fitted standard deviations (SDs) of the multivariate log-GARCH(1,1) and log-GARCH(7,1)-X* models, and the nominal and relative differences between the SDs (computed as log-GARCH(1,1) minus log-GARCH(7,1)-X* and log-GARCH(1,1) over log-GARCH(7,1)-X*, respectively)